

FORM PTO-150 (Modified)  
(REV 11-99)

U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

ATTORNEY'S DOCKET NUMBER

TRANSMITTAL LETTER TO THE UNITED STATES  
DESIGNATED/ELECTED OFFICE (DO/EO/US)  
CONCERNING A FILING UNDER 35 U.S.C. 371

6433/80968

Pat Sel

U.S. APPLICATION NO. (IF KNOWN, SEE 37 CFR

09/701132

INTERNATIONAL APPLICATION NO.

PCT/AU99/00385

INTERNATIONAL FILING DATE

21 MAY 1999

PRIORITY DATE CLAIMED

21 MAY 1998

O I P E

TITLE OF INVENTION

ANTIGENS AND THEIR DETECTION

APPLICANT(S) FOR DO/EO/US

Peter Richard Reeves and Lei Wang



Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☐ This is an express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).
4. ☐ A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. ☒ A copy of the International Application as filed (35 U.S.C. 371 (c) (2))
  - a. ☒ is transmitted herewith (required only if not transmitted by the International Bureau).
  - b. ☒ has been transmitted by the International Bureau.
  - c. ☐ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☐ A translation of the International Application into English (35 U.S.C. 371(c)(2)).
7. ☐ A copy of the International Search Report (PCT/ISA/210).
8. ☒ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371 (c)(3))
  - a. ☒ are transmitted herewith (required only if not transmitted by the International Bureau).
  - b. ☒ have been transmitted by the International Bureau.
  - c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
  - d. ☐ have not been made and will not be made.
9. ☐ A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
10. ☐ An oath or declaration of the inventor(s) (35 U.S.C. 371 (c)(4)).
11. ☐ A copy of the International Preliminary Examination Report (PCT/IPEA/409).
12. ☐ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371 (c)(5)).

Items 13 to 20 below concern document(s) or information included:

13. ☐ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
14. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
15. ☒ A **FIRST** preliminary amendment.
16. ☐ A **SECOND** or **SUBSEQUENT** preliminary amendment.
17. ☐ A substitute specification.
18. ☐ A change of power of attorney and/or address letter.
19. ☐ Certificate of Mailing by Express Mail
20. ☐ Other items or information:

U.S. APPLICATION NO. (IF KNOWN, SEE 37 CFR

INTERNATIONAL APPLICATION NO.

ATTORNEY'S DOCKET NUMBER

097701132

PCT/AU99/00385

6433/80968

21. The following fees are submitted:

**BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) - (5)) :**

- ☐ Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO ..... **\$1,000.00**
- ☐ International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO ..... **\$860.00**
- ☐ International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO ..... **\$710.00**
- ☐ International preliminary examination fee paid to USPTO (37 CFR 1.482) but all claims did not satisfy provisions of PCT Article 33(1)-(4) ..... **\$690.00**
- ☐ International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4) ..... **\$100.00**

**ENTER APPROPRIATE BASIC FEE AMOUNT =****\$1,000.00**Surcharge of **\$130.00** for furnishing the oath or declaration later than months from the earliest claimed priority date (37 CFR 1.492 (e)). ☐ 20 ☐ 30**\$0.00**

CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE	
Total claims	31 - 20 =	11	x \$18.00	<b>\$198.00</b>
Independent claims	2 - 3 =	0	x \$80.00	<b>\$0.00</b>
Multiple Dependent Claims (check if applicable).			<input type="checkbox"/>	<b>\$0.00</b>
<b>TOTAL OF ABOVE CALCULATIONS =</b>				<b>\$1,198.00</b>

Reduction of 1/2 for filing by small entity, if applicable. Verified Small Entity Statement must also be filed (Note 37 CFR 1.9, 1.27, 1.28) (check if applicable).

☒**\$599.00****SUBTOTAL =****\$599.00**Processing fee of **\$130.00** for furnishing the English translation later than months from the earliest claimed priority date (37 CFR 1.492 (f)). ☐ 20 ☐ 30**\$0.00****TOTAL NATIONAL FEE =****\$599.00**

Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31) (check if applicable).

☐**\$0.00****TOTAL FEES ENCLOSED =****\$599.00**

Amount to be refunded	\$
charged	\$

☒ A check in the amount of **\$599.00** to cover the above fees is enclosed.

☐ Please charge my Deposit Account No. \_\_\_\_\_ in the amount of \_\_\_\_\_ to cover the above fees.  
A duplicate copy of this sheet is enclosed.

☒ The Commissioner is hereby authorized to charge any fees which may be required, or credit any overpayment to Deposit Account No. **23-0920** A duplicate copy of this sheet is enclosed.

**NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.**

SEND ALL CORRESPONDENCE TO:

SIGNATURE

Shannon L. Nebolsky, Esq.

NAME

41,217

REGISTRATION NUMBER

November 20, 2000

DATE

## IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: Peter Richard Reeves	)	
and Lei Wang	)	Attorney Docket:
	)	6433/80968
U.S. Serial No.: Not yet assigned	)	
	)	
Filed: November 21, 2000	)	
	)	Examining Group:
For: ANTIGENS AND THEIR	)	1600
DETECTION	)	
	)	
Examiner: Not yet assigned	)	

PRELIMINARY AMENDMENT and SEQUENCE LISTING

Commissioner for Patents  
Washington, D.C. 20231

Sir:

The subject application is a U.S. National Phase filing under 35 U.S.C. 371 based on International Patent Application Serial No. PCT/AU99/00385, international filing date May 21, 1999, claiming the benefit of foreign priority filing of Australian Patent Application Serial No. PP 3634, filed May 21, 1998.

ABSTRACT:

After the claims, please insert the following Abstract of the invention.

--ABSTRACT

The invention relates to novel nucleotide sequences located in a gene which encodes a bacterial flagellin antigen, and the use of those nucleotide sequences for the detection of bacteria which express particular flagellin antigens, on the basis of that antigen alone, or in conjunction with the O antigen expressed by that strain.--

IN THE CLAIMS:

Please amend the claims as follows:

21. A method according to [any one of claims 8, 9, 11, 15 or 19] claim 7 wherein the sample is selected from the group consisting of a sample derived from food, a sample derived from faeces and a sample derived from a patient or animal.

22. A kit for identifying the H serotype of *E. coli*, the kit comprising at least one nucleic acid molecule according to [any one of claims 1 to 6] claim 1.

23. A kit for identifying the H and O serotype of *E. coli*, the kit comprising:

(a) at least one nucleic acid molecule according to [any one of claims 1 to 6] claim 1; and

(b) at least one nucleic acid molecule derived from and specific for a gene encoding a transferase or a gene encoding an enzyme for the transport or



processing of a polysaccharide or oligosaccharide unit, the gene being involved in the synthesis of a particular *E. coli* O antigen.

Please insert the following new claims.

26. A method according to claim 9 wherein the sample is selected from the group consisting of a sample derived from food, a sample derived from faeces and a sample derived from a patient or animal.

27. A method according to claim 11 wherein the sample is selected from the group consisting of a sample derived from food, a sample derived from faeces and a sample derived from a patient or animal.

28. A method according to claim 15 wherein the sample is selected from the group consisting of a sample derived from food, a sample derived from faeces and a sample derived from a patient or animal.

29. A method according to claim 19 wherein the sample is selected from the group consisting of a sample derived from food, a sample derived from faeces and a sample derived from a patient or animal.

30. A kit for identifying the H serotype of *E. coli*, the kit comprising at least one nucleic acid molecule according to claim 6.

31. A kit for identifying the H and O serotype of *E. coli*, the kit comprising:

(a) at least one nucleic acid molecule according to claim 6; and

(b) at least one nucleic acid molecule derived from and specific for a gene encoding a transferase or a gene encoding an enzyme for the transport or processing of a polysaccharide or oligosaccharide unit, the gene being involved in the synthesis of a particular *E. coli* O antigen.

#### REMARKS

##### I. The Amendments

Claims 21 through 23 were amended to remove multiple dependency. New claims 26 to 31 were added to maintain the claimed subject matter as filed before removal of multiple dependency. The new claims are supported by the claims originally filed. A typographical error in claim dependency was also corrected in claim 21. No new matter has been added to the subject patent application by virtue of this preliminary amendments.

Claims 1 through 31 are in the case and are before the Examiner. An Abstract page has been added to the specification. The text for the abstract was taken from the PCT application abstract, thereby adding no new matter.

Pursuant to the new rules of practice in patent cases before the U.S. Patent and Trademark Office, a clean copy of the claims before the Examiner after entry of the Preliminary Amendment are enclosed. For the convenience of the Patent Office, a clean copy of the new Abstract page is also enclosed.

II. Biological Sequence Listing  
Statements under 1.825(a) and 1.821(f)

This Preliminary Amendment including Statements under 37 C.F.R. 1.825(a) and 1.821(f) is accompanied by substitute sheets for the paper copy of the Sequence Listing of the above-identified patent application. The content of the Sequence Listing is the same as that of the Sequence Listing for the international application as filed in the PCT, the difference being that the format has been updated in the paper copy to conform to the current U.S. Patent Office Sequence Listing requirements with page numbers beginning at 1.

This paper is also accompanied by a write-protected diskette (3.50 inch, 1.44 Mb storage capacity) containing the computer readable form (CRF) of the Sequence Listing as ASCII output from PatentIn version 2.0. The computer readable form filename is

"P30384.app". The CRF of the sequence listing was generated by the PCT-filing associate in Australia using Patentin Version 2.0 on May 21, 1999 on a PC-compatible computer.

The Patentin output was transmitted via e-mail and copied onto the enclosed diskette November 21, 2000 unaltered as received. The information recorded in the computer readable form is identical to the enclosed paper copy of the Sequence Listing. A copy of the Patentin output was opened into a word processing program separately to produce the enclosed paper copy substitute sheets of the Biological Sequence Listing that has the appropriate page numbering. The substitute sheets include no new matter.

#### SUMMARY


The claims and specification have been preliminarily amended to conform to U.S. practice, and substitute pages incorporating the preliminary amendments are enclosed, along with a computer-readable form of the Biological Sequence Listing to supplement the paper copy already transmitted from the international authority.

The application is believed to be in condition for allowance. An early notice to that effect is earnestly solicited.

No further fee or petition is believed to be necessary. However, should any further fee be needed, please charge our Deposit Account No. 23-0920, and deem this paper to be the required petition.

The Examiner is requested to phone the undersigned should any questions arise that can be dealt with over the phone to expedite this prosecution.

Respectfully submitted,

  
Shannon L. Nebolsky, Reg. No. 41,217

Enclosures:

Diskette with file P30384.app

Welsh & Katz, Ltd.  
120 South Riverside Plaza  
22nd Floor  
Chicago, Illinois 60606  
(312) 655-1500

CERTIFICATE OF EXPRESS MAILING

I hereby certify that this Preliminary Amendment and Sequence Listing, together with the stated enclosures, is being deposited with the United States Postal Service as Express Mail (EL617904151US) in an envelope addressed to: Commissioner for Patents, Washington, D.C. 20231 on November 21, 2000.

Slavko J. Petrović

CLAIMS

1. A nucleic acid molecule which encodes all or part of an *E. coli* flagellin protein, the molecule being capable of identifying the H serotype of an *E. coli* when hybridised to a gene of the *E. coli* which encodes a flagellin protein, provided that the molecule does not encode a flagellin protein expressed by the *E. coli* H1, H7, H12 or H48 type strains.

2. A nucleic acid molecule according to claim 1 wherein the molecule is derived from a *fliC* gene.

3. A nucleic acid molecule according to claim 1 including all or part of a sequence according to any one of SEQ ID NOS:1 to 68.

4. A nucleic acid molecule according to claim 1 consisting of all or part of a sequence according to any one of SEQ ID NOS: 1 to 68.

5. A nucleic acid molecule according to claim 4 wherein the molecule is from about 10 to 20 nucleotides in length.

6. A primer selected from the group of primers shown in Table 3.

7. A method of detecting the H serotype of *E. coli* in a sample, the method comprising the following steps:

(a) contacting a gene of an *E. coli* in the sample with a nucleic acid molecule according to claim 1 in conditions sufficient to allow the nucleic acid molecule to hybridise to the gene; and

(b) detecting a nucleic acid molecule which is hybridised to the gene, to detect the H serotype of the *E. coli* in the sample.

8. A method according to claim 7 wherein the hybridised nucleic acid molecules are detected by Southern Blot analysis.

9. A method of detecting the H serotype of *E. coli* in a sample, the method comprising the following steps:

(a) contacting a gene of an *E. coli* in the sample with a pair of nucleic acid molecules according to claim 1 in conditions sufficient to allow the pair of nucleic acid molecules to hybridise to the gene; and

(b) detecting a pair of nucleic acid molecules which is hybridised to the gene, to detect the H serotype of the *E. coli* in the sample.

10. A method according to claim 9 wherein the hybridised pairs of nucleic acid molecules are detected by the polymerase chain reaction.

11. A method for detecting the H and O serotype of *E. coli* in a sample, the method comprising the following steps:



(a) contacting a gene of the *E. coli* with a nucleic acid molecule derived from a gene encoding a transferase or a gene encoding an enzyme for the transport or processing of a polysaccharide or oligosaccharide unit, the gene being involved in the synthesis of a *E. coli* O antigen, in conditions sufficient to allow the nucleic acid molecule to hybridise to the gene;

(b) contacting a gene of an *E. coli* in the sample with a nucleic acid molecule according to claim 1 in conditions sufficient to allow the nucleic acid molecule to hybridise to the gene; and

(c) detecting nucleic acid molecules which are hybridised to the genes, to detect the H and O serotype of the *E. coli* in the sample.

12. A method according to claim 11 wherein the nucleic acid molecule of step (a) is selected from the group consisting of:

*wbdH* (nucleotide position 739 to 1932 of Figure 5),  
*wzx* (nucleotide position 8646 to 9911 of Figure 5),  
*wzy* (nucleotide position 9901 to 10953 of Figure 5),  
*wbdM* (nucleotide position 11821 to 12945 of Figure 5),  
*wbdN* (nucleotide position 79 to 861 of Figure 6),  
*wbdO* (nucleotide position 2011 to 2757 of Figure 6),  
*wbdP* (nucleotide position 5257 to 6471 of Figure 6),  
*wbdR* (nucleotide position 13156 to 13821 of Figure 6),  
*wzx* (nucleotide position 2744 to 4135 of Figure 6) and  
*wzy* (nucleotide position 858 to 2042 of Figure 6).

13. A method according to claim 12 wherein the nucleic acid molecule of step (a) is a primer

selected from the group of primers shown in Tables 8, 8A, 9 and 9A.

14. A method according to claim 11 wherein the hybridised nucleic acid molecules are detected by Southern Blot analysis.

15. A method for detecting the H and O serotype of *E. coli* in a sample, the method comprising the following steps:

(a) contacting a gene of the *E. coli* with a pair of nucleic acid molecules derived from a gene encoding a transferase or a gene encoding an enzyme for the transport or processing of a polysaccharide or oligosaccharide unit, the gene being involved in the synthesis of a *E. coli* O antigen, in conditions sufficient to allow the pair of nucleic acid molecules to hybridise to the gene;

(b) contacting a gene of an *E. coli* in the sample with a pair of nucleic acid molecules according to claim 1 in conditions sufficient to allow the pair of nucleic acid molecules to hybridise to the gene; and

(c) detecting pairs of nucleic acid molecules which are hybridised to the genes, to detect the H and O serotype of the *E. coli* in the sample.

16. A method according to claim 15 wherein the pair of nucleic acid molecules of step (a) is selected from the group consisting of:

*wbdH* (nucleotide position 739 to 1932 of Figure 5),

*wzx* (nucleotide position 8646 to 9911 of Figure 5),

wzy (nucleotide position 9901 to 10953 of Figure 5),  
wbdM (nucleotide position 11821 to 12945 of Figure 5),  
wbdN (nucleotide position 79 to 861 of Figure 6),  
wbdO (nucleotide position 2011 to 2757 of Figure 6),  
wbdP (nucleotide position 5257 to 6471 of Figure 6),  
wbdR (nucleotide position 13156 to 13821 of Figure 6),  
wzx (nucleotide position 2744 to 4135 of Figure 6) and  
wzy (nucleotide position 858 to 2042 of Figure 6).

17. A method according to claim 15 wherein the nucleic acid molecules of the pair of step (a) are primers selected from the group of primers shown in Tables 8, 8A, 9 and 9A.

18. A method according to claim 15 wherein the hybridised pairs of nucleic acid molecules are detected by the polymerase chain reaction.

19. A method for detecting the H and O serotype of *E. coli* in a sample, the method comprising the following steps:

(a) contacting a gene of an *E. coli* in the sample with a nucleic acid molecule according to claim 1, in conditions sufficient to allow the nucleic acid molecule to hybridise to the gene; and

(b) detecting a nucleic acid molecule which is hybridised to the gene, to detect the H and O serotype of *E. coli* in the sample.

20. A method according to claim 19 wherein the nucleic acid molecule is according to any one of SEQ ID NOS: 9, 55, 57 to 65.

21. A method according to claim 7 wherein the sample is selected from the group consisting of a sample derived from food, a sample derived from faeces and a sample derived from a patient or animal.

22. A kit for identifying the H serotype of *E. coli*, the kit comprising at least one nucleic acid molecule according to claim 1.

23. A kit for identifying the H and O serotype of *E. coli*, the kit comprising:

(a) at least one nucleic acid molecule according to claim 1; and

(b) at least one nucleic acid molecule derived from and specific for a gene encoding a transferase or a gene encoding an enzyme for the transport or processing of a polysaccharide or oligosaccharide unit, the gene being involved in the synthesis of a particular *E. coli* O antigen.

24. A kit according to claim 23 wherein the at least one nucleic acid molecule of (a) is selected from the group consisting of:

*wbdH* (nucleotide position 739 to 1932 of Figure 5),  
*wzx* (nucleotide position 8646 to 9911 of Figure 5),  
*wzy* (nucleotide position 9901 to 10953 of Figure 5),  
*wbdM* (nucleotide position 11821 to 12945 of Figure 5),  
*wbdN* (nucleotide position 79 to 861 of Figure 6),  
*wbdO* (nucleotide position 2011 to 2757 of Figure 6),  
*wbdP* (nucleotide position 5257 to 6471 of Figure 6),

*wbdR* (nucleotide position 13156 to 13821 of Figure 6),  
*wzx* (nucleotide position 2744 to 4135 of Figure 6) and  
*wzy* (nucleotide position 858 to 2042 of Figure 6).

25. A kit according to claim 24 wherein the nucleic acid molecule of (a) is a primer selected from the group of primers shown in Tables 8, 8A, 9 and 9A.

26. A method according to claim 9 wherein the sample is selected from the group consisting of a sample derived from food, a sample derived from faeces and a sample derived from a patient or animal.

27. A method according to claim 11 wherein the sample is selected from the group consisting of a sample derived from food, a sample derived from faeces and a sample derived from a patient or animal.

28. A method according to claim 15 wherein the sample is selected from the group consisting of a sample derived from food, a sample derived from faeces and a sample derived from a patient or animal.

29. A method according to claim 19 wherein the sample is selected from the group consisting of a sample derived from food, a sample derived from faeces and a sample derived from a patient or animal.

30. A kit for identifying the H serotype of *E. coli*, the kit comprising at least one nucleic acid molecule according to claim 6.

31. A kit for identifying the H and O serotype of *E. coli*, the kit comprising:

(a) at least one nucleic acid molecule according to claim 6; and

(b) at least one nucleic acid molecule derived from and specific for a gene encoding a transferase or a gene encoding an enzyme for the transport or processing of a polysaccharide or oligosaccharide unit, the gene being involved in the synthesis of a particular *E. coli* O antigen.

ABSTRACT

The invention relates to novel nucleotide sequences located in a gene which encodes a bacterial flagellin antigen, and the use of those nucleotide sequences for the detection of bacteria which express particular flagellin antigens, on the basis of that antigen alone, or in conjunction with the O antigen expressed by that strain.

Antigens and Their DetectionTECHNICAL FIELD

The invention relates to novel nucleotide sequences located in a gene which encodes a bacterial flagellin antigen, and the use of those nucleotide sequences for the detection of bacteria which express particular flagellin antigens, on the basis of that antigen alone, or in conjunction with the O antigen expressed by that strain.

BACKGROUND ART

The flagellum of many bacteria appears to be made up of a single protein known as flagellin. The serotyping schemes of *E. coli* and *Salmonella enterica* are based on highly variable antigenic surface structures which include the lipopolysaccharide which carries the O antigen and flagellin which is now known to be the carrier of the classical H antigen. In many strains of *S. enterica* there are two loci (*fliC* and *fljB*) which encode flagellin, and a regulatory system which allows one only to be expressed at any time; and which also provides for expression to rapidly alternate between the two forms first identified as two phases (H1 and H2) for the H antigen of most strains. In *E. coli* there are 54 forms of H antigen recognised and until recently they were all thought to be encoded at the *fliC* locus, as has been shown for *E. coli* K-12. However in the 1980s Ratiner [Ratiner Y A "Phase variation of the H antigen in *Escherichia coli* strain B1327-41, the standard strain for *Escherichia coli* flagellin antigen H3" FEMS Microbiol. Lett 15 (1982) 33-36; Ratiner Y A "Presence of two structural genes determining antigenically different phase-specific flagellins in some *Escherichia coli* strains" FEMS Microbiol. Lett. 19 (1983) 37-41; Ratiner Y A "Two genetic arrangements determining flagellin antigen specificities in two diphasic *Escherichia coli* strains" FEMS Microbiol. Lett. 29 (1985) 317-323; Ratiner Y A "Different alleles of the flagellin gene *hagB* in *Escherichia coli* standard H



test strains" FEMS Microbiol Lett. 48 (1987) 97-104.] showed that in some cases there are two loci and that expression can alternate. The matter was further complicated by a recent paper by Ratiner [Ratiner Y A (1998) "New flagellin-specifying genes in some *Escherichia coli* strains" J. Bacteriol. 180 979-984] showing three loci (*flk*, *flI* and *flm*) for flagellin in addition to *fliC* although the *fljB* locus has not been found in *E. coli*. However *E. coli* strains are normally identified by the combination of one O antigen and one H antigen [and K antigen when present as a capsule (K) antigen], with no problems reported for the vast majority of cases with alternate phases, while *S. enterica* strains are normally identified by the combination of O, H1 and H2 antigens. It is still not clear how widespread in *E. coli* H antigens determined by flagellin genes other than *fliC* are.

Typing is typically carried out using specific antisera. The incidence of pathogenic *E. coli* in association with human and animal disease supports the need for suitable and rapid typing techniques.

#### DESCRIPTION OF THE INVENTION

In a first aspect, the present invention provides a novel nucleic acid molecule encoding all or part of an *E. coli* flagellin protein.

The present invention provides, for the first time, full length sequence for a flagellin gene for the following *E. coli* type strains: H6 (SEQ ID NO: 8), H9 (SEQ ID NO: 11), H10 (SEQ ID NO: 12), H14 (SEQ ID NO: 15), H18 (SEQ ID NO: 18), H23 (SEQ ID NO: 22), H51 (SEQ ID NO: 50), H45 (SEQ ID NO: 43), H49 (SEQ ID NO: 48), H19 (SEQ ID NO: 19), H30 (SEQ ID NO: 29), H32 (SEQ ID NO: 31), H26 (SEQ ID NO: 25), H41 (SEQ ID NO: 39), H15 (SEQ ID NO: 16), H20 (SEQ ID NO: 20), H28 (SEQ ID NO: 27), H46 (SEQ ID NO: 44), H31 (SEQ ID NO: 30), H34 (SEQ ID NO: 33), H43 (SEQ ID NO: 41) and H52 (SEQ ID NO: 51). Corrected full length sequences have been obtained for H7 (SEQ ID NO: 9) and

- 3 -

H12(SEQ ID NO: 14) type strains.

Partial flagellin gene sequence, including the central variable region, has been obtained for the following *E. coli* H type strains: H40(SEQ ID NO: 38), H8(SEQ ID NO: 10), H21(SEQ ID NO: 21), H47(SEQ ID NO: 46), H11(SEQ ID NO: 13), H17(SEQ ID NO: 17), H25(SEQ ID NO: 24), H42(SEQ ID NO: 40), H27(SEQ ID NO: 26), H35(SEQ ID NO: 34), H2(SEQ ID NO: 67), H3(SEQ ID NO: 68), H24(SEQ ID NO: 23), H37(SEQ ID NO: 35), H50(SEQ ID NO: 49), H4(SEQ ID NO: 6), H44(SEQ ID NO: 42), H38(SEQ ID NO: 36), H39(SEQ ID NO: 37), H55(SEQ ID NO: 53), H29(SEQ ID NO: 28), H33(SEQ ID NO: 32), H5(SEQ ID NO: 7), H54(SEQ ID NO: 52) and H56(SEQ ID NO: 54).

Comparison of sequences demonstrates that unique flagellin genes have now been sequenced (partially or completely) for the following *E. coli* H type strains: H1, H2, H3, H5, H6, H7, H9, H11, H12, H14, H15, H18, H19, H20, H21, H23, H24, H25, H26, H27, H28, H29, H30, H31, H32, H33, H34, H35, H37, H38, H39, H41, H42, H43, H45, H46, H48, H49, H51, H52, H54, and H56 and either H8 or H40, H10 or H50 and H4 or H17.

By comparison of these sequences, the present inventors were able to identify specific sequences for each of the above H serotypes.

The present invention also provides *fliC* sequences from 10 different H7 strains, in addition to that from the H7 type strain, and two sequences specific to H7 of O157 and O55 *E. coli* strains.

The present invention encompasses all or part of the flagellin genes sequenced for H2, H3, H5, H6, H9, H11, H14, H18, H19, H20, H21, H23, H24, H25, H26, H27, H28, H29, H30, H31, H32, H33, H34, H35, H37, H38, H39, H41, H42, H43, H44, H45, H46, H47, H48, H49, H51, H52, H54, H55, H56, H8, H40, H15, H10, or H50, H4 and H17 type strains. Of these flagellin genes sequenced, those from the type strains for H8 and H40 are identical, those from type strains H10 and H50, H1 and H12, H38 and H55, H21 and

H47, and H4, H17 and H44 type strains are highly similar.

The invention also encompasses newly provided sequence for H7 and H12 as well as novel primers for the specific amplification of H1, H7, H12 and H48 as well as  
5 for the other above mentioned newly sequenced flagellin genes.

By cloning and expression of these sequenced flagellin genes in a *fliC* deletion *E. coli* K-12 strain, and use of anti-H antiserum, we have confirmed the H  
10 specificities encoded by 39 flagellin genes. The 39 H specificities are H1, H2, H4, H5, H6, H7, H9, H10, H11, H12, H14, H15, H16, H18, H19, H20, H21, H23, H24, H26, H27, H28, H29, H30, H31, H32, H33, H34, H38, H39, H41, H42, H43, H45, H46, H49, H51, H52, and H56, encoded by  
15 flagellin genes obtained from H type strains for H1, H2, H4, H5, H6, H7, H9, H10, H11, H12, H14, H15, H3, H18, H19, H20, H21, H23, H24, H26, H27, H28, H29, H30, H31, H32, H33, H34, H38, H39, H41, H42, H43, H45, H46, H49, H51, H52, and H56 respectively.

The nucleic acid molecules of the invention may be variable in length. In one embodiment they are oligonucleotides of from about 10 to about 20 nucleotides  
20 in length. The oligonucleotides of the invention are specific for the flagellin gene from which they are derived and are derived from the central region of the gene. In one embodiment, oligonucleotides in accordance with the present invention, which also include  
25 oligonucleotides from the previously sequenced *E. coli* H1, H7, H12 and H48 genes, are those shown in Table 3.

The 45 sequences (see Table 3) provide a panel to which newly sequenced genes can be compared to select  
30 specific oligonucleotides for those newly sequenced genes.

In a second aspect the invention provides a method of detecting the presence of *E. coli* of a particular H  
35 serotype in a sample, the method comprising the step of specifically hybridising at least one nucleic acid molecule derived from a flagellin gene, wherein the at

least one nucleic acid molecule is specific for a particular flagellin gene associated with the H serotype, to any *E. coli* in the sample which contain the gene, and detecting any specifically hybridised nucleic acid molecules, wherein the presence of specifically hybridised nucleic acid molecules identifies the presence of the H serotype in the sample.

In one preferred embodiment the detection method is a Southern blot method. More preferably, the nucleic acid molecule is labelled and hybridisation of the nucleic acid molecule is detected by autoradiography or detection of fluorescence.

Preferred nucleic acid molecules for the detection of particular flagellin genes are listed in Table 3.

In a third aspect the invention provides a method of detecting the presence of *E. coli* of a particular H serotype in a sample, the method comprising the step of specifically hybridising at least one pair of nucleic acid molecules to any *E. coli* in the sample which contains the flagellin gene for the particular H serotype, wherein at least one of the nucleic acid molecules is specific for the particular flagellin gene associated with the H serotype, and detecting any specifically hybridised nucleic acid molecules, wherein the presence of specifically hybridised nucleic acid molecules identifies the presence of the H serotype in the sample.

In one preferred embodiment the detection method is a polymerase chain reaction method. More preferably, the nucleic acid molecules are labelled and hybridisation of the nucleic acid molecule is detected by electrophoresis.

It is recognised that there may be instances where spurious hybridisation will arise through the initial selection of a sequence found in many different genes but this is typically recognisable by, for instance, comparison of band sizes against controls in PCR gels, and an alternative sequence can be selected.

In a fourth aspect the invention provides a method for detecting the presence of a particular O serotype and H serotype of *E. coli* in a sample, the method comprising the following steps:

5 (a) specifically hybridising at least one nucleic acid molecule, derived from and specific for a gene encoding a transferase or a gene encoding an enzyme for the transport or processing of a polysaccharide or oligosaccharide unit, the gene being involved in the  
10 synthesis of a particular *E. coli* O antigen, to any *E. coli* in the sample which contain the gene;

(b) specifically hybridising at least one nucleic acid molecule derived from and specific for a particular flagellin gene associated with that H serotype, to any *E. coli* in the sample which contain the gene; and  
15

(c) detecting any specifically hybridised nucleic acid molecules.

Preferred nucleic acid molecules for the detection of particular flagellin genes are listed in Table 3.

20 In one preferred embodiment, the sequence of the nucleic acid molecule specific for the O antigen is specific to the nucleotide sequence encoding the O111 antigen. More preferably, the sequence is derived from a gene selected from the group consisting of *wbdH*  
25 (nucleotide position 739 to 1932 of Figure 5), *wzx* (nucleotide position 8646 to 9911 of Figure 5), *wzy* (nucleotide position 9901 to 10953 of Figure 5), *wbdM* (nucleotide position 11821 to 12945 of Figure 5) and fragments of those molecules of at least 10-12 nucleotides  
30 in length. Particularly preferred nucleic acid molecules are those set out in Tables 8 and 8A, with respect to the above mentioned genes.

In another preferred embodiment, the sequence of the nucleic acid molecule specific for the O antigen is  
35 specific to the nucleotide sequence encoding the O157 antigen. More preferably, the sequence is derived from a gene selected from the group consisting of *wbdN*

(nucleotide position 79 to 861 of Figure 6), *wbdO*  
(nucleotide position 2011 to 2757 of Figure 6), *wbdP*  
(nucleotide position 5257 to 6471 of Figure 6), *wbdR*  
(nucleotide position 13156 to 13821 of Figure 6), *wzx*  
5 (nucleotide position 2744 to 4135 of Figure 6) and *wzy*  
(nucleotide position 858 to 2042 of Figure 6) and  
fragments of those molecules of at least 10-12 nucleotides  
in length. Particularly preferred nucleic acid molecules  
are those set out in Tables 9 and 9A, with respect to the  
10 above mentioned genes.

In one preferred embodiment the detection method is a  
Southern blot method. More preferably, the nucleic acid  
molecule is labelled and hybridisation of the nucleic acid  
molecule is detected by autoradiography or detection of  
15 fluorescence.

In a fifth aspect the invention provides a method for  
detecting the presence of a particular O serotype and H  
serotype of *E. coli* in a sample, the method comprising the  
following steps:

20 (a) specifically hybridising at least one pair of  
nucleic acid molecules, at least one of which is derived  
from and specific for a gene encoding a transferase or a  
gene encoding an enzyme for the transport or processing of  
a polysaccharide or oligosaccharide unit, the gene being  
25 involved in the synthesis of the particular *E. coli* O  
antigen, to any *E. coli* in the sample which contain the  
gene;

(b) specifically hybridising at least one pair of  
nucleic acid molecules, at least one of which is derived  
30 from and specific for a particular flagellin gene  
associated with the particular H serotype, to any *E. coli*  
in the sample which contain the gene; and

(c) detecting any specifically hybridised nucleic  
acid molecules.

35 Preferred nucleic acid molecules for the detection of  
particular flagellin genes are listed in Table 3.

In one preferred embodiment, the sequence of the nucleic acid molecule specific for the O antigen is specific to the nucleotide sequence encoding the O111 antigen. More preferably, the sequence is derived from a gene selected from the group consisting of *wbdH* (nucleotide position 739 to 1932 of Figure 5), *wzx* (nucleotide position 8646 to 9911 of Figure 5), *wzy* (nucleotide position 9901 to 10953 of Figure 5), *wbdM* (nucleotide position 11821 to 12945 of Figure 5) and fragments of those molecules of at least 10-12 nucleotides in length. Particularly preferred nucleic acid molecules are those set out in Tables 8 and 8A, with respect to the above mentioned genes.

In another preferred embodiment, the sequence of the nucleic acid molecule specific for the O antigen is specific to the nucleotide sequence encoding the O157 antigen. More preferably, the sequence is derived from a gene selected from the group consisting of *wbdN* (nucleotide position 79 to 861 of Figure 6), *wbdO* (nucleotide position 2011 to 2757 of Figure 6), *wbdP* (nucleotide position 5257 to 6471 of Figure 6), *wbdR* (nucleotide position 13156 to 13821 of Figure 6), *wzx* (nucleotide position 2744 to 4135 of Figure 6) and *wzy* (nucleotide position 858 to 2042 of Figure 6) and fragments of those molecules of at least 10-12 nucleotides in length. Particularly preferred nucleic acid molecules are those set out in Tables 9 and 9A, with respect to the above mentioned genes.

In one preferred embodiment the detection method is a polymerase chain reaction method. More preferably, the nucleic acid molecules are labelled and hybridisation of the nucleic acid molecule is detected by electrophoresis.

The present inventors believe that based on the teachings of the present invention and available information concerning O antigen gene clusters, and through use of experimental analysis, comparison of nucleic acid sequences or predicted protein structures, nucleic acid molecules in accordance with the invention

can be readily derived for any particular O antigen of interest. Suitable bacterial strains can typically be acquired commercially from depositary institutions.

There are currently 166 defined *E. coli* O antigens.

5        Samples of the 166 different *E. coli* O antigen serotypes are available from Statens Serum Institut, Copenhagen, Denmark.

10        The inventors envisage rare circumstances whereby two genetically similar gene clusters encoding serologically different O antigens have arisen through recombination of genes or mutation so as to generate polymorphic variants.

15        In these circumstances multiple pairs of oligonucleotides may be selected to provide hybridisation to the specific combination of genes. The invention thus envisages the use of a panel containing multiple nucleic acid molecules for use in the method of testing for O antigen in conjunction with H antigen, wherein the nucleic acid molecules are derived from genes encoding transferases and/or enzymes for the transport or processing of a polysaccharide or oligosaccharide unit including *wzx* or *wzy* genes, wherein the panel of nucleic acid molecules is specific to a particular O antigen. The panel of nucleic acid molecules can include nucleic acid molecules derived from O antigen sugar pathway genes where necessary.

25        The inventors also found two mutated flagellin genes from H type strains for H35 and H54 which have insertion sequences inserted into normal flagellar genes identical or near identical to that of the H11 and H21 type strains respectively. Thus, primers for H11 and H21 (listed in Table 3) would also amplify fragments in H35 and H54, which differ in sizes to those in H11 and H21 respectively. The inventors also provide two pairs of primers each for H35 and H54 based on the insertion sequence (see H35 and H54 columns in Table 3). The use of  
30        one of them in combination with one of the H11 or H21 primers will generate a PCR band only in H35 or H54 respectively, and this will also differentiate H35 and H54  
35



from H11 and H21 respectively.

The present invention also relates to methods of detecting the presence of particular *E. coli* H antigens or H antigen and O antigen combinations where one or more nucleic acid molecules which generate a particular size fragment indicative of the presence of that H antigen are used or in which the combination of one antigen specific primer for that H antigen with another primer for a related H antigen provides for the detection of the particular H antigen by hybridisation to the relevant gene. Preferably, the H antigen is H11, H21, H35 or H54.

The pairs of nucleic acid molecules where the method of the fifth aspect is used may both hybridise to the relevant H or O antigen gene or alternatively only one may hybridise to the relevant gene and the other to another site.

The inventors recognise in applying the methods of the invention for detecting combinations of O and H antigens to samples, that the methods do not indicate whether a positive result for a particular O and H antigen combination arises because the O and H antigen are present on a single *E. coli* strain present in the sample or are present on different *E. coli* strains present in the sample. Because the ability to identify the presence of *E. coli* strains with particular O and H antigen combinations is highly desirable (due to the relationship between particular combinations and pathogenicity) the determination that a particular combination is present in a sample can be followed by isolation of single colonies and checking whether they contain the relevant combination by using the same method again or using antibody labelled magnetic beads to separate cells expressing the particular O or H antigen and then testing the isolated cells for the other serotype.

In addition, as mentioned above, the present inventors have established the existence of H7 primers specific to the O157 and O55 serotypes. Using such

- 11 -

primers it is possible to detect particular O and H antigen combinations with the use of H specific nucleic acid molecules.

5 In a sixth aspect the invention provides a method for detecting the presence of a particular O serotype and H serotype of *E. coli* in a sample, the method comprising the following steps:

10 (a) specifically hybridising at least one nucleic acid molecule, derived from and specific for a gene encoding a flagellin associated with a particular *E. coli* H antigen serotype to any *E. coli* carrying the gene and present in the sample;  
and

15 (b) detecting the at least one specifically hybridised nucleic acid molecule, wherein the at least one nucleic acid molecule is specific for the particular combination of O and H antigen.

Preferably the combination is O55:H7 or O157:H7.

20 The ability to detect the O157:H7 combination from a particular H7 primer or pair is of particular use given the association of this combination with pathogenic strains.

25 In a seventh aspect the present invention provides a method for testing a food derived sample for the presence of one or more particular *E. coli* O antigens and H antigens comprising testing the sample by a method of the fourth, fifth or sixth aspect the invention.

30 In an eighth aspect the present invention provides a method for testing a faecal derived sample for the presence of one or more particular *E. coli* O antigens and H antigens comprising testing the sample by a method of the fourth, fifth or sixth aspect the invention.

35 In a ninth aspect the present invention provides a method for testing a patient or animal derived sample for the presence of one or more particular *E. coli* O antigens and H antigens comprising testing the sample by a method of the fourth, fifth or sixth aspect the invention.

Preferably, the method of the seventh, eighth or ninth aspect of the invention is a polymerase chain reaction method. More preferably the oligonucleotide molecules for use in the method are labelled. Even more preferably the hybridised nucleic acid molecules are detected by electrophoresis.

In the above described methods it will be understood that where pairs of nucleic acid molecules are used one of the nucleic acid molecules may hybridise to a sequence that is not from the O antigen transferase, wzx or wzy gene or the flagellin gene. Further where both hybridise to these genes the O antigen molecules may hybridise to the same or a different one of these genes.

In a tenth aspect the present invention provides a kit for identifying the H serotype of *E. coli*, the kit comprising:

at least one nucleic acid molecule derived from and specific for an *E. coli* flagellin gene.

In an eleventh aspect the present invention provides a kit for identifying the H and O serotype of *E. coli*, the kit comprising:

(a) at least one nucleic acid molecule derived from and specific for an *E. coli* flagellin gene; and

(b) at least one nucleic acid molecule derived from and specific for a gene encoding a transferase or a gene encoding an enzyme for the transport or processing of a polysaccharide or oligosaccharide unit, the gene being involved in the synthesis of a particular *E. coli* O antigen.

The nucleic acid molecules may be provided in the same or different vials. The kit may also provide in the same or separate vials a second set of specific nucleic acid molecules.

Particularly preferred nucleic acid molecules for inclusion in the kits are those specified in Tables 3, 8, 8A, 9 and 9A as described above.

## DEFINITIONS

In this specification, we have used term "flagellin gene" in many cases where previously one would have used "*fliC*", to allow for the uncertainty as to locus introduced by recent observations. However, uncertainty as to the locus does not alter the fact that most *E. coli* strains express a single H antigen and that a single flagellin gene sequence per strain is required to give the genetic basis for H antigen variation. Any use of the name *fliC* in this specification where a different locus is later shown to be involved would not affect the validity of conclusions drawn regarding application of information based on the sequence, where the conclusions do not relate to the map position. Thus it is generally the nucleic acid molecule itself which is of importance rather than the name attributed to the gene. When it is known or suspected that the gene encoding the H antigen is not in the *fliC* locus, we use the term flagellin rather than *fliC*.

The phrase, "a nucleic acid molecule derived from a gene" means that the nucleic acid molecule has a nucleotide sequence which is either identical or substantially similar to all or part of the identified gene. Thus a nucleic acid molecule derived from a gene can be a molecule which is isolated from the identified gene by physical separation from that gene, or a molecule which is artificially synthesised and has a nucleotide sequence which is either identical to or substantially similar to all or part of the identified gene. While some workers consider only the DNA strand with the same sequence as the mRNA transcribed from the gene, here either strand is intended.

Transferase genes are regions of nucleic acid which have a nucleotide sequence which encodes gene products that transfer monomeric sugar units.

Flippase or *wzx* genes are regions of nucleic acid which have a nucleotide sequence which encodes a gene

product that flips oligosaccharide repeat units generally composed of three to six monomeric sugar units to the external surface of the membrane.

Polymerase or *wzy* genes are regions of nucleic acid which have a nucleotide sequence which encodes gene products that polymerise repeating oligosaccharide units generally composed of 3-6 monomeric sugar units.

The nucleotide sequences provided in this specification are described as anti-sense sequences. This term is used in the same manner as it is used in Glossary of Biochemistry and Molecular Biology Revised Edition, David M. Glick, 1997 Portland Press Ltd., London on page 11 where the term is described as referring to one of the two strands of double-stranded DNA usually that which has the same sequence as the mRNA. We use it to describe this strand which has the same sequence as the mRNA.

#### NOMENCLATURE

##### Synonyms for *E. coli* O111 *rfb*

Current names	Our names	Bastin et al. 1991
wbdH	orf1	
gmd	orf2	
wbdI	orf3	orf3.4*
manC	orf4	rfbM*
manB	orf5	rfbK*
wbdJ	orf6	orf6.7*
wbdK	orf7	orf7.7*
wzx	orf8	orf8.9 and rfbX*
wzy	orf9	
wbdL	orf10	
wbdM	orf11	

\* Nomenclature according to Bastin D.A., et al. 1991 "Molecular cloning and expression in *Escherichia coli* K-12 of the *rfb* gene cluster determining the O antigen of an *E. coli* O111 strain". Mol. Microbiol. 5:9 2223-2231.

##### Other Synonyms

wzy	rfc
wzx	rfbX
rmlA	rfbA
rmlB	rfbB
rmlC	rfbC
rmlD	rfbD
glf	orf6*
wbbI	orf3#, orf8* of <i>E. coli</i> K-12

wbbJ orf2#, orf9\* of *E. coli* K-12  
 wbbK orf1#, orf10\* of *E. coli* K-12  
 wbbL orf5#, orf 11\* of *E. coli* K-12  
 # Nomenclature according to Yao, Z. And M. A. Valvano 1994.

- 5 "Genetic analysis of the O-specific lipopolysaccharide biosynthesis region (rfb) of *Escherichia coli* K-12 W3110: identification of genes the confer groups-specificity to *Shigella flexneri* serotypes Y and 4a". *J. Bacteriol.* 176: 4133-4143.
- \* Nomenclature according to Stevenson et al. 1994. "Structure of the O-antigen of *E. coli* K-12 and the sequence of its rfb gene cluster". *J. Bacteriol* 176: 4144-4156.
- 10 • The O antigen genes of many species were given rfb names (rfbA etc) and the O antigen gene cluster was often referred to as the rfb cluster. There are now new names for the rfb genes as shown
- 15 in the table. Both terminologies have been used herein, depending on the source of the information.

In the claims that follow and in the summary of the invention, except where the context requires otherwise due to express language or necessary implication, the word "comprising" is used in the sense of "including", i.e. the features specified may be associated with further features in various embodiments of the invention.

#### BRIEF DESCRIPTION OF THE DRAWINGS

- 25 Figure 1 shows *Eco* RI restriction maps of cosmid clones pPR1054, pPR1055, pPR1056, pPR1058, pPR1287 which are subclones of *E. coli* O111 O antigen gene cluster. The thickened line is the region common to all clones. Broken lines show segments that are non-contiguous on the chromosome. The deduced restriction map for *E. coli* strain M92 is shown above.
- 30

- Figure 2 shows a restriction mapping analysis of *E. coli* O111 O antigen gene cluster within the cosmid clone pPR1058. Restriction enzymes are: (B: *Bam*HI; Bg: *Bgl*II, E: *Eco*RI; H: *Hind*III; K: *Kpn*I; P: *Pst*I; S: *Sal*I and X: *Xho*I. Plasmids pPR1230, pPR1231, and pPR1288 are deletion derivatives of pPR1058. Plasmids pPR 1237, pPR1238, pPR1239 and pPR1240 are in pUC19. Plasmids pPR1243, pPR1244, pPR1245, pPR1246 and pPR1248 are in pUC18, and pPR1292 is in pUC19. Plasmid pPR1270 is in
- 40

- 16 -

pT7T319U. Probes 1, 2 and 3 were isolated as internal fragments of pPR1246, pPR1243 and pPR1237 respectively. Dotted lines indicate that subclone DNA extends to the left of the map into attached vector.

5           Figure 3 shows the structure of *E. coli* O111 O antigen gene cluster.

Figure 4 shows the structure of *E. coli* O157 O antigen gene cluster.

10           Figure 5 shows the nucleotide sequence (SEQ ID NO: 45) of the *E. coli* O111 O antigen gene cluster. Note: (1) The first and last three bases of a gene are underlined and of italic respectively.; (2) The region which was previously sequenced by Bastin and Reeves 1995 "Sequence and anlysis of the O antigen gene (rfb) cluster of *Escherichia coli* O111" Gene 164: 17-23 is marked.

15           Figure 6 shows the nucleotide sequence (SEQ ID NO: 56) of the *E. coli* O157 O antigen gene cluster. Note: (1) The first and last three bases of a gene (region) are underlined and of italic respectively (2) The region previously sequenced by Bilge et al. 1996 "Role of the *Escherichia coli* O157-H7 O side chain in adherence and analysis of an rfb locus". Inf. and Immun 64:4795-4801 is marked.

20           Figures 7 to 9 show the nucleotide sequences (SEQ ID NOS: 66 to 68 respectively) obtained for flagellin genes from *E. coli* type strains for H1 to H3 respectively. The primer positions listed in Table 3 are based on treating the first nucleotide of each of these sequences as No. 1.

25           Figures 10 to 18 show the nucleotide sequences (SEQ ID NOS: 6 to 14 respectively) obtained for flagellin genes from *E. coli* type strains for H4 to H12 respectively. The primer positions listed in Table 3 are based on treating the first nucleotide of each of these sequences as No. 1.

30           Figures 19 and 20 show the nucleotide sequences (SEQ ID NOS: 15 to 16 respectively) obtained for flagellin genes from *E. coli* type strains for H14 and H15 respectively. The primer positions listed in Table 3 are

based on treating the first nucleotide of each of these sequences as No. 1.

5        Figures 22 and 26 show the nucleotide sequences (SEQ ID NOS: 17 to 21 respectively) obtained for flagellin genes from *E. coli* type strains for H17 and H21 respectively. The primer positions listed in Table 3 are based on treating the first nucleotide of each of these sequences as No. 1.

10       Figures 27 to 39 show the nucleotide sequences (SEQ ID NOS: 22 to 34) obtained for flagellin genes from *E. coli* type strains for H23 to H35 respectively. The primer positions listed in Table 3 are based on treating the first nucleotide of each of these sequences as No. 1.

15       Figures 40 to 49 show the nucleotide sequences (SEQ ID NOS: 35 to 44) obtained for flagellin genes from *E. coli* type strains for H37 to H46 respectively. The primer positions listed in Table 3 are based on treating the first nucleotide of each of these sequences as No. 1.

20       Figures 50 to 55 show the nucleotide sequences (SEQ ID NOS: 46 to 51) obtained for flagellin genes from *E. coli* type strains for H47 to H52 respectively. The primer positions listed in Table 3 are based on treating the first nucleotide of each of these sequences as No. 1.

25       Figures 56 to 58 show the nucleotide sequences (SEQ ID NOS: 52 to 54) obtained for flagellin genes from *E. coli* type strains for H54 to H56 respectively. The primer positions listed in Table 3 are based on treating the first nucleotide of each of these sequences as No. 1.

30       Figure 59 shows the nucleotide sequence (SEQ ID NO: 55) obtained for the flagellin gene from *E. coli* H7 strain M1179. The primer positions listed in Table 3 are based on treating the first nucleotide of each of these sequences as No. 1.

35       Figures 60 to 68 show the nucleotide sequences (SEQ ID NOS: 57 to 65 respectively) obtained for flagellin genes from *E. coli* strains M1004, M1211, M1200, M1686, M1328, M917, M527, M973, and M918 respectively. The primer



positions listed in Table 3 are based on treating the first nucleotide of each of these sequences as No. 1.

Figure 69 shows the nucleotide sequence (SEQ ID NO: 1) of the *fliC* gene and DNA flanking the *fliC* gene from the H25 type strain.

Figure 70A shows the nucleotide sequence (SEQ ID NO: 2) obtained from the 5' end of the insert of plasmid pPR1989. The insert of plasmid pPR1989 encodes the second flagellin gene of the H55 type strain.

Figure 70B shows the nucleotide sequence (SEQ ID NO: 3) obtained from the 3' end of the insert of plasmid pPR1989. The insert of plasmid pPR1989 encodes the second flagellin gene of the H55 type strain.

Figure 71 shows the nucleotide sequence (SEQ ID NO:4) obtained from the 5' end of the insert of plasmid pPR1993. The insert of plasmid pPR1993 encodes the second flagellin gene of the H36 strain.

Figure 72 shows the nucleotide sequence (SEQ ID NO:5) obtained from the 3' end of the insert of plasmid pPR1993. The insert of plasmid pPR1993 encodes the second flagellin gene of the H36 type strain.

Figure 73 A shows the sequence of polylinker and the SD sequence of plasmid pTrc99A.

Figure 73B shows the sequence of the junction region between the SD sequence and the start of flagellin gene in the plasmids used for the expression of flagellin genes.

#### BEST METHOD OF CARRYING OUT THE INVENTION

In carrying out the methods of the invention with respect to the testing of particular sample types including samples from food, patients, animals and faeces the samples are prepared by routine techniques routinely used in the preparation of such samples for DNA based testing. The steps for testing the samples using particular nucleic acid molecules in assay formats such as Southern blots and PCR are performed under routinely determined conditions appropriate to the sample and the

nucleic acid molecules.

## H antigen

### Materials and Methods

#### 5 1. Bacterial strains and plasmid:

There are 54 H types in *E. coli* [Ewing, W.H.: Edwards and Ewing's identification of the *Enterobacteriaceae*, Elsevier Science Publishers, Amsterdam, The Netherlands, 1986]: note H antigens from 1 to 57 were listed and that  
10 13, 22 and 57 are not valid. All the standard H type strains except H16 were obtained from the Institute of Medical and Veterinary Science, Adelaide, Australia. The primary stocks are held at the Statens Serum Institut, Copenhagen, Denmark.

15 The additional H7 strains used are listed in Table 1.

We do not have the type strain for H16. It is known that the H3 type strain is biphasic and can also express the H16 flagellin gene [Ratiner, Y. A. (1985) "Two genetic arrangements determining flagellar antigen specificities  
20 in two diphasic *E. coli* strains. FEMS Microbiol Lett 19: 317-323]. We have sequenced and cloned the H16 flagellin gene from the H3 type strain (see below).

*E. coli* K-12 strain C600 *hsm hsr fliC::Tn10*  
[Kuwajima, G. (1988) "Flagellin domain that affects H  
25 antigenicity of *E. coli* K-12" J. Bacteriol. 170; 485-488] (laboratory stock no. M2126) was obtained from Dr Benita Westerlund-Wikstrom of the Department of Biosciences, University of Helsinki, Finland. *E. coli* K-12 strain EJ2282 (laboratory no. P5560) is a *fliC* deletion strain,  
30 and was obtained from Dr Masatoshi Enomoto of the Department of Biology, Okayama University, Japan [Tominaga, A. M. A.-H. Mahmoud, T. Mokaiharu and M. Enomoto (1994) "Molecular characterization of intact but cryptic, flagellin genes in the genus *Shigella*." Mol.  
35 Microbiol. 12: 277-285].

Plasmid pTrc99A was purchased from Pharmacia LKB (Melbourne, VIC, Australia).

## 2. Antisera

Antisera against H1, H3, H8, H14, H15, H17, H23, H24, H25, H26, H29, H30, H31, H32, H33, H35, H36, H37, H38, H39, H43, H44, H46, H47, H48, H49, H52, H53, H54, H55, and H56 were obtained from the Institute of Medical and Veterinary Science, Adelaide, Australia. Antisera against H2, H4, H5, H6, H7, H9, H10, H11, H12, H16, H18, H19, H20, H21, H27, H28, H34, H40, H41, H42, H45, and H51 were obtained from Denka Seiken Co., Ltd, Tokyo, Japan.

Antisera to type H50 was not available from any known source.

The antisera available were checked against the appropriate type strains to confirm the specificities of both flagellin H antigen and H antisera: 52 sera (all those except anti-H16 serum listed above) gave a positive reaction with the corresponding type strains for that serum.

## 3. Agglutination test:

Bacteria from 1 ml of an overnight culture grown in Luria broth (Difco Tryptone, 10g/l; Difco yeast extract, 5g/l; NaCl, 0.5 g/l; pH 7.2) at 30°C was centrifuged (4000 rpm/10 min) and the bacteria pellet resuspended in 100 ml of saline. The agglutination test was carried out by mixing equal volumes (5 ml) of both the cells and antiserum on a slide. The slide was rocked for 1 minute and then observed for agglutination. For all agglutination tests, saline containing no antiserum was mixed with cells to be used as a negative control.

For testing the H specificities of strain M2126 or strain P5560 carrying plasmid containing cloned flagellin genes, cells of M2126 or P5560 were used as an additional negative control.

All agglutination tests were first carried out using undiluted antisera (note that the antisera we used have been diluted before reaching our hands), except for anti-

H11, anti-H34, anti-H52 and anti-H26 serum for which we used 1:10 dilutions to avoid background agglutination. In cases for which cross-reactions have been reported, we carried out agglutination tests using serial dilutions of sera (see section 10.1)

#### 4. Motility test:

The motility of strain M2126 or strain P5560 carrying cloned flagellin genes was examined microscopically. 1 ml of overnight culture grown in Luria broth (Difco Tryptone, 10g/l; Difco yeast extract, 5g/l; NaCl, 0.5 g/l; pH 7.2) at 30°C was inoculated into 10 ml of Luria broth, and the culture was shaken at 100 rpm at 30°C to early log phase (OD 625 = 0.2). A loopful of culture was placed on a slide and examined under a microscope. Motility of individual cells was easily distinguished from Brownian movement and streaming, and presence or absence of motility recorded.

#### 5. Isolation of chromosomal DNA:

Chromosomal DNA from all the 53 H type strains and the strains listed in Table 1 was isolated using the Promega Genomic isolation kit (Madison WI USA). Each chromosomal DNA sample was checked by gel electrophoresis of the DNA and by PCR amplification of the *mdh* gene using oligonucleotides based on the *E. coli* K-12 *mdh* gene [Boyd, E.F., Nelson, K., Wang, F.-S., Whittam, T.S. and Selander, R.K.: Molecular genetic basis of allelic polymorphism in malate dehydrogenase (*mdh*) in natural populations of *Escherichia coli* and *Salmonella enterica*. Proc. Natl. Acad. Sci. USA 91 (1994) 1280-1284].

#### 6. PCR amplification of flagellin gene:

Flagellin genes from different strains were first PCR amplified using one of the following four pairs of oligonucleotides:

#1285 (5'-atggcacaagtcattaatac) and  
#1286 (5'-ttaaccctgcagtagagaca);

- 22 -

#1417 (5'-ctgatcactcaaaataatatcaac) and  
#1418 (5'-ctgcggtacctggttggc);  
#1431 (5'-atggcacaagtcattaataaccaac) and  
#1432 (5'-ctaaccctgcagcagagaca):  
5 #1575 (5'-gggtggaaaccaataacg) and  
#1576(5'-gcgcatacaggcaatttgg)

PCR reactions were carried out under the following conditions: denaturing, 94°C/30'; annealing, temperature varies (refer to Table 2)/30'; extension, 72°C/1'; 30  
10 cycles. The PCR product was purified using the Promega Wizard PCR purification kit (Madison WI USA) before being sequenced.

The H36 and H53 type strains gave two PCR bands using  
15 primer pairs #1431/#1432 and #1417/#1418 respectively, and were not sequenced.

#### 7. Enzymes and buffers:

Restriction endonucleases and DNA T4 ligase were purchased from Boehringer Mannheim (Castle Hill, NSW, Australia). Restriction enzymes were used in the  
20 recommended commercial buffer.

#### 8. Sequencing of the flagellin genes:

Each PCR product was first sequenced using the  
25 oligonucleotide primers used for the PCR amplification. Primers based on the obtained sequence were then used to sequence further, and this procedure was repeated until the entire PCR product was sequenced.

The sequencing reactions were performed using the  
30 DyeDeoxy Terminator Cycle Sequencing method (Applied Biosystems, CA, USA), and reaction products were analysed using fluorescent dye and an ABI377 automated sequencer (CA, USA).

Sequence data were processed and analysed using  
35 Staden programs [Sacchi CT, Zanella R C, Caugant D A, Frasc C E, Hidalgo N T, Milagres L G, Pessoa L L, Ramos S R, Camargo M C C and Melles C E A "Emergence of a new

- 23 -

- clone of serogroup C *Neisseria meningitidis* in Sao Paulo, Brazil" J. Clin. Microbiol. 30 (1992) 1282-1286;
- Staden, R.: Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. Nucl. Acids Res. 10 (1982a) 4731-4751;
- 5 Staden, R.: An interactive graphics program for comparing and aligning nucleic acid and amino acid sequences. Nucl. Acids Res. 10 (1982b) 2951-2961;
- Staden, R.: Computer methods to locate signals in nucleic acid sequences. Nucl. Acids Res. 12 (1984a) 505-519;
- 10 Staden, R.: Graphic methods to determine the function of nucleic acid sequences. A summary of ANALYSEQ options. Nucl. Acids Res. 12 (1984b) 521-538;
- Staden, R.: The current status and portability of our sequence handling software. Nucl. Acids Res. 14 (1986) 217-231).
- 15

We were able to PCR amplify flagellin genes from H type strains for H7, 23, 12, 51, 45, 49, 19, 9, 30, 32, 26, 41, 15, 20, 28, 46, 31, 14, 18, 6, 34, 48, 43, 10, 52, and also from H7 strains m1004, m527, m1686, m1211, m1328, m973, m1179, m1200, m917, and m918 using primers #1575 and #1576 which are based on sequences 51-34 bp upstream and 37-54 bp downstream of start and end of the *E. coli* K-12 *fliC* gene respectively. Thus, the full sequence of the flagellin gene from these strains was obtained and the use of flanking sequence for primers makes it highly likely that they are at the *fliC* locus.

20

25

For other strains, we were only able to amplify the flagellin gene using one or more of the other three pairs of primers, which are based on sequence within the *fliC* gene, and thus only partial sequence was obtained. These amplicons may be of the *fliC* gene or one of the alternative flagellin genes. The flagellin gene sequences from H type strains for H40, 8, 21, 47, 11, 27, 35, 2, 3, 24, 37, 50, 4, 44, 38, 55, 29, 33, 5, and 56 obtained are lacking 18 and 14 codons at 5' and 3' ends respectively. The flagellin gene sequence of H39 obtained using primers

30

35

#1285/#1286 lacks 18 and 19 codons at 5' and 3' ends respectively. The flagellin gene sequence of H type strains of H17, 25 and 42 lack 23 and 21 codons at 5' and 3' ends respectively. The flagellin gene sequence of the H type strain for H54 lacks 23 and 12 codons at the 5' and 3' ends respectively. There is very little variation in the sequence at the two ends of flagellin genes and antigenic variation is due to variation in the central region of the gene. The absence of sequence for the ends of some of the flagellin genes is not important for the purpose of the present invention relating to the detection of antigenic variation by DNA sequence based means.

The *fliC* genes from H type strains of H1, H7 and H12 have been sequenced previously [Schoenhals, G. and Whitfield, C.: Comparative analysis of flagellin sequences from *Escherichia coli* strains possessing serologically distinct flagellar filaments with a shared complex surface pattern. J. Bacteriol. 175 (1993) 5395-5402] and we did not sequence the gene from the H1 strain.

We have sequenced *fliC* genes from a set of H7 strains with different O antigens, including that of *fliC* from the H7 type strain as one of the set: we have found four differences from the published H7 sequence (GenBank accession number L07388) which we believe are due to errors in the published sequence.

We have also re-sequenced the *fliC* gene from the H12 type strain, and have found one difference from the published H12 sequence (GenBank accession number L07389) which we believe is due to an error in the published sequence.

The flagellin genes from type strains H35 and H54 were also amplified using primers #1431/#1432, which are based on sequence within the *fliC* gene. Sequence data revealed that these two genes would be non-functional due to insertion sequence inserted in the middle of them. We have sequenced them to facilitate selection of primers for the functional flagellin genes.

### 9. Cloning of flagellin genes

DNA was digested for 2 hr at 37°C with appropriate restriction enzyme(s). The reaction product was then extracted once with phenol, and twice with ether. DNA was precipitated with 2 vols of ethanol and resuspended in water before the ligation reaction was carried out. Ligation was carried out O/N at 4°C and the ligated DNA was electroporated into one of the *E. coli fliC* mutant strains.

9.1. Cloning of flagellin genes from type strains for H1, H2, H3, H5, H6, H7, H9, H10, H11, H12, H14, H15, H18, H19, H20, H21, H24, H26, H27, H28, H29, H31, H34, H38, H39, H41, H42, H43, H45, H46, H49, H51, H52, and H56:

The full flagellin gene was PCR amplified using primers #1868 and #1870 (Table 3A). Both these primers are based on the sequences of the H7 flagellin gene of the H7 type strain. #1868 is the 5' primer: there is an *NcoI* site incorporated into the primer (Table 3B) and the flagellin gene starts at base 3 of the *NcoI* site. The 3' primer #1870 has a *BamHI* site incorporated downstream of the stop codon of the flagellin gene (Table 3B). PCR reactions were carried out under the following conditions: denaturing, 94°C/30'; annealing, temperature varies (refer to Table 3A)/30'; extension, 72°C/1'; 30 cycles. The PCR product was purified using the Promega Wizard PCR purification kit (Madison WI USA) before being digested by restriction enzymes *NcoI* and *BamHI* and cloned into the *NcoI/BamHI* sites of plasmid pTrc99A.

Plasmid pTrc99A has a strong *trc* promoter upstream of the polylinker. Downstream of the promoter, it contains the ribosome binding site (SD sequence, see Fig 73) which is located 8bp upstream of the ATG site within the *NcoI* site. The polylinker and the SD sequence of pTrc99A are shown in Fig 73.

The plasmids generated were given pPR numbers, and



- 26 -

they are listed in Table 3A. In these plasmids, the expression module consists of the *trc* promoter, the SD sequence, and the full flagellin gene. The sequence of the junction region between the SD sequence and the start of flagellin gene is shown in Fig 73.

For flagellin genes from type strains for H6, H7, H9, H10, H12, H14, H18, H19, H20, H26, H28, H31, H41, H43, H45, H46, H49, H51, and H52, we have the full sequence for each gene and the primer sequences (#1868 and #1870) are conserved among them. The cloned genes therefore have the same sequence as those from the type strains.

For flagellin genes from type strains for H1, H15 and H34, we also have the full sequence. The previously published sequence of the flagellin gene from the H1 type strain was extracted from GenBank (accession number L07387) and used. Primer #1868 is conserved in all three. But, primer #1870 has the third base of the fifth last codon in the H1 sequence changed from A to G, and the third base of the second last codon changed from C to T in the H15 and H34 sequences: these changes did not change the amino acid coded, so the cloned genes encode the same gene products as those of the type strains.

For flagellin genes from type strains for H2, H3, H5, H11, H21, H24, H27, H29, H38, H39, H42, and H56, we do not have the full sequences. In the plasmids carrying genes from these type strains, the expression module consists of the *trc* promoter, the SD sequence, and the full flagellin gene with the first and the last 21 base pairs being determined by the primer sequences which are based on the H7 flagellin gene of the H7 type strain. The sequence of the junction region between the SD sequence and the start of flagellin gene is shown in Fig 73.

## 9.2. Cloning of the flagellin gene from type strain of H23:

The full flagellin gene was PCR amplified using primers #1868 and #1869 (Table 3A). #1868 is the 5'

primer: there is an *NcoI* site incorporated into the primer (Table 3B) and the flagellin gene starts at base 3 of the *NcoI* site. The 3' primer #1869 has a *SalI* site incorporated downstream of the stop codon of the flagellin gene (Table 3B). PCR reactions were carried out under the following conditions: denaturing, 94°C/30'; annealing, 55°C/30'; extension, 72°C/1'; 30 cycles. The PCR product was purified using the Promega Wizard PCR purification kit (Madison WI USA) before being digested by restriction enzymes *NcoI* and *SalI* and cloned into the *NcoI/SalI* sites of plasmid pTrc99A to give plasmid pPRI942.

Plasmid pTrc99A has a strong *trc* promoter upstream of the polylinker. Downstream of the promoter, it contains the ribosome binding site (SD sequence, see Fig 73) which is located 8bp upstream of the ATG site within the *NcoI* site. The polylinker and the SD sequence of pTrc99A are shown in Fig 73.

The expression module of pPRI942 consists of the *trc* promoter, the SD sequence, and the full flagellin gene. The sequence of the junction region between the SD sequence and the start of flagellin gene is shown in Fig 73.

### 9.3. Cloning of flagellin genes from type strains of H30, H32 and H33:

The full flagellin gene was PCR amplified using primers #1868 and #1871 (Table 3A). #1868 is the 5' primer: there is an *NcoI* site incorporated into the primer (Table 3B) and the flagellin gene starts at base 3 of the *NcoI* site. The 3' primer #1871 has a *PstI* site incorporated downstream of the stop codon of the flagellin gene (Table 3B). PCR reactions were carried out under the following conditions: denaturing, 94°C/30'; annealing, temperature varies (refer to Table 3A)/30'; extension, 72°C/1'; 30 cycles. The PCR product was purified using the Promega Wizard PCR purification kit (Madison WI USA) before being digested by restriction enzymes *NcoI* and *PstI*

and cloned into the *Nco*I/*Pst*I sites of plasmid pTrc99A.

Plasmid pTrc99A has a strong *trc* promoter upstream of the polylinker. Downstream of the promoter, it contains the ribosome binding site (SD sequence, see Fig 73) which is located 8bp upstream of the ATG site within the *Nco*I site. The polylinker and the SD sequence of pTrc99A are shown in Fig 73.

For flagellin genes from type strains for H30 and H32, we have the full sequence. Primer #1868 sequence is conserved in both of them. But, primer #1871 has the third base of the fourth last codon in both sequences changed from G to A to remove a *Pst*I site (see Table 3B): this change did not change the amino acid coded. The expression module consists of the *trc* promoter, the SD sequence, and the full flagellin gene coding for a gene product which is same as that of the type strain. The sequence of the junction region between the SD sequence and the start of flagellin gene is shown in Fig 73.

We do not have the full sequence for the flagellin gene from the H33 type strain. In the plasmid containing the H33 type strain gene, the expression module consists of the *trc* promoter, the SD sequence, and the full flagellin gene with the first and the last 21 base pairs been determined by the primer sequences which were used for the cloning of H30 and H32. The sequence of the junction region between the SD and the start of flagellin gene is shown in Fig 73.

#### 9.4. Flagellin genes from type strains for H4 and H17:

For the flagellin genes of H4 and H17 type strains the full sequence was not obtained, and the sequenced parts were PCR amplified and cloned into plasmid pPR1951 to give in each case a gene in which the first 26 and the last 31 codons are based on the sequence of the H7 flagellin gene of the H7 type strain.

##### 9.4.1

##### Construction of expression plasmid vector

*pPR1951:*

The first 26 codons of the H7 flagellin gene was first PCR amplified using primers #1868 and #1872 (Table 3B). #1868 is the 5' primer: there is an *Nco*I site incorporated into the primer (Table 3B) and the flagellin gene starts at base 3 of the *Nco*I site. Primer #1872 was made to have the last two codons (codons 25 and 26) changed from CTG TCG (Leucine and Serine) to GGA TCC (Glycine and Serine) to generate a *Bam*HI site. This PCR fragment was digested with *Nco*I and *Bam*HI before being cloned into the *Nco*I/*Bam*HI sites of pTrc99A to make plasmid pPR1949.

The last 31 codons (including the stop codon) of the H7 flagellin gene was PCR amplified using primers #1884 and #1871 (Table 3A). The 5' primer, #1884, has the first two of the 31 codons changed from TCG AAA (Serine and Lysine) to TCT AGA (Serine and Arginine) to generate a *Xba*I site (Table 3B). The 3' primer #1871 has a *Pst*I site incorporated downstream of the stop codon (Table 3B). This PCR fragment was digested with *Xba*I and *Pst*I, and then cloned into the *Xba*I/*Pst*I sites of pPR1949 to make plasmid pPR1951.

#### 9.4.2 Cloning of flagellin genes from the H4 and H17 type strains:

The central regions of flagellin genes from type strains H4 and H17 were PCR amplified using primers #1878 and #1885 (Table 3B), which have a *Bam*HI and a *Xba*I incorporated at their ends respectively. PCR reactions were carried out under the following conditions: denaturing, 94°C/30'; annealing, 65°C/30'; extension, 72°C/1'; 30 cycles. The PCR product was purified using the Promega Wizard PCR purification kit (Madison WI USA) before being digested by restriction enzymes *Bam*HI and *Xba*I and cloned into the *Xba*I/*Bam*HI sites of plasmid pPR1951 to make plasmids pPR1955 (H4) and pPR1957 (H17).

The expression module of plasmids pPR1955 and pPR1957

consists of the *trc* promoter, the SD sequence, the first 24 codons of the H7 flagellin gene (of the H7 type strain), 2 codons encoding Glycine and Serine, 292 or 293 codons of the central region based on the flagellin gene obtained from the H4 or H17 type strain respectively, 2 codons encoding Serine and Arginine, and then the last 29 codons of the H7 flagellin gene (of the H7 type strain).

10. Expression of flagellin gene plasmids in *E. coli* strains lacking the *fliC* gene, and identification of the H antigens encoded by these plasmids:

Plasmids carrying flagellin genes as described in section 9 (see Table 3A for a list) were electroporated into strains M2126 or P5560. Strains M2126 and P5560 do not have functional *fliC* genes, and are not motile when examined under a microscope. Transformants carrying any of the plasmids listed in Table 3A are motile when examined under a microscope. Thus, the flagellin genes in all of the plasmids are expressed.

The antigenic specificity of the flagellin of each transformant was then determined by slide agglutination.

10.1 Flagellin genes from type strains for H2, H5, H6, H7, H9, H11, H14, H15, H18, H19, H20, H21, H23, H24, H26, H27, H28, H29, H30, H31, H32, H33, H34, H39, H41, H42, H43, H45, H46, H49, H51, H52, and H56:

As shown in Table 3A, strains with plasmids carrying these flagellin genes expressed the same H antigen as their respective flagellin parent strain.

For flagellin specificities H2, H5, H6, H7, H9, H14, H15, H18, H19, H20, H23, H24, H26, H27, H28, H29, H31, H33, H39, H51, H52, and H56, there was no cross reaction reported between these flagellins and flagellin antisera for other H antigens [Ewing, W. H.: Edwards and Ewing's identification of the *Enterobacteriaceae*, Elsevier Science Publishers, Amsterdam, The Netherlands, 1986], and we conclude that we have in each case sequenced the gene

encoding the flagellin of the expected specificity from the respective type strain.

It has been observed that cross reactions exist between some type strains and certain antisera at different levels of dilution (of the antisera), being H11 with anti-H21 and anti-H40, H21 with anti-H11, H30 with anti-H32, H32 with anti-H30, H34 with anti-H24 and anti-H31, H41 with anti-H37 and anti-H39, H42 with anti-H6, H43 with anti-H37, H45 with anti-H20, H46 with anti-H17, and H49 with anti-H39 [Ewing, W. H.: Edwards and Ewing's identification of the *Enterobacteriaceae*, Elsevier Science Publishers, Amsterdam, The Netherlands, 1986]. We have tested strain M2126 or strain P5560 carrying plasmids containing flagellin genes obtained from each of these type strains (H11, H21, H30, H32, H34, H41, H42, H43, H45, H46, and H49) with the appropriate cross-reacting antisera.

For strain M2126 or strain P5560 carrying plasmids containing flagellin genes obtained from type strains H11, H34, H41, H42, H43, H45, H46, and H49, no cross reaction was found. We conclude that we have in each case sequenced the gene coding the flagellin of the expected specificity from the respective type strain.

Cross reaction was observed for strain P5560 carrying plasmid pPR1948 (containing the flagellin gene obtained from the H30 type strain) with anti-H32 serum, strain P5560 carrying pPR1940 (containing the flagellin gene obtained from the H32 type strain) with anti-H30 serum, and strain M2126 carrying plasmid pPR1995 (containing the flagellin gene obtained from the H21 type strain) with anti-H11 serum.

We note that the reported cross reactions between the H30 type strain and anti-H32, the H32 type strain and anti-H30, and the H21 type strain and anti-H11 happened at a higher level of dilution (of antisera) than for all other type strains with the antisera mentioned above [Ewing, W. H.: Edwards and Ewing's identification of the

*Enterobacteriaceae*., Elsevier Science Publishers, Amsterdam, The Netherlands, 1986]. We conclude that except for these three cases, the antiserum used were supplied at a dilution which did not exhibit cross reactions. For the  
5 three strains carrying flagellin genes cloned from type strains for H30, H32, and H21, it was necessary to further dilute the antiserum.

Strain P5560 carrying plasmid pPR1948 (containing the flagellin gene obtained from the H30 type strain)  
10 agglutinates with anti-H30 serum when the antiserum is diluted to 1:60, but agglutinates with anti-H32 serum only at a dilution of 1:10 and not at a 1:20 dilution (note that the antisera we used have been diluted before reaching our hands). In contrast, strain P5560 carrying  
15 plasmid pPR1940 (containing flagellin gene obtained from the H32 type strain) agglutinates with anti-H32 serum when the antiserum is diluted at 1:100, but agglutinates with anti-H30 serum only at a 1:10 dilution and not at a 1:10 dilution. Thus, we conclude that the flagellin genes we  
20 sequenced from type strains for H30 and H32 encode flagellins of H30 and H32 specificities respectively.

Strain M2126 carrying plasmid pPR1995 (containing the flagellin gene obtained from the H21 type strain)  
25 agglutinates with anti-H21 serum when the antiserum is diluted to 1:40, but agglutinates only with undiluted anti-H11 serum and not at a 1:10 dilution (note that the antisera we used have been diluted before reaching our hands). In contrast, strain M2126 carrying plasmid pPR1981  
30 (containing flagellin gene obtained from the H11 type strain) did not agglutinate with anti-H21 serum. Thus, we conclude that the flagellin genes we sequenced from type strains for H21 encodes flagellin of H21 specificity.

#### 10.2 Flagellin genes from type strains of H1 and 35 H12:

These two genes are very similar in sequence, with 8 a.a difference between the gene products. It has been

known that some cross-reaction exists between anti-H12 serum and the H1 type strain and between anti-H1 serum and the H12 type strain [Ewing, W. H.: Edwards and Ewing's identification of the *Enterobacteriaceae*, Elsevier Science Publishers, Amsterdam, The Netherlands, 1986]. Strain M2126 carrying pPR1920 (carrying a flagellin gene from the H1 type strain, Table 3A) agglutinates with anti-H1 serum when the antiserum is diluted to 1:100, but agglutinates only with undiluted anti-H12 serum and not at a 1:10 dilution (please note that the antisera we used have been diluted before reaching our hands). In contrast, strain M2126 carrying plasmid pPR1990 (carrying a flagellin gene from the H12 type strain, Table 3A) agglutinates with anti-H12 serum when the antiserum is diluted at 1:100, but agglutinates only with undiluted anti-H1 serum and not at a 1:10 dilution. Thus, we conclude that the flagellin genes we sequenced from type strains for H1 and H12 encode flagellins of H1 and H12 specificities respectively.

#### 10.3. Flagellin gene coding for H16:

Strain P5560 carrying plasmid pPR1969 agglutinated with anti-H16 serum. pPR1969 carries a flagellin gene amplified from the H3 type strain. It has been shown that this H3 type strain is a biphasic strain which can express H3 and H16 specificities [Ratiner, Y. A. (1985) "Two genetic arrangements determining flagellar antigen specificities in two diphasic *E. coli* strains. FEMS Microbiol Lett 19: 317-323]. Thus, the H3 type strain has two flagellin genes coding for H3 and H16 specificities. We conclude that we have cloned and sequenced the H16 flagellin gene from this H3 type strain.

#### 10.4 Flagellin gene coding for H4:

The flagellin genes obtained from type strains for H4 and H17 are nearly identical, with 4 a.a. difference in the gene products. Plasmid pPR1955 carries a flagellin



gene from the H4 type strain, and plasmid pPR1957 carries a flagellin gene from the H17 type strain. Strain P5560 carrying plasmid pPR1955 or plasmid pPR1957 agglutinated with anti-H4 serum, but not with anti-H17 serum. It has been shown that the type strain for H17 is a biphasic strain which can express H17 and H4 [Ratiner, Y. A. (1985) "Two genetic arrangements determining flagellar antigen specificities in two diphasic *E. coli* strains. FEMS Microbiol Lett 19: 317-323]. The flagellin gene obtained from type strain for H44 is also highly similar to that obtained from the H4 type strain, with 2 a.a. difference in the gene products. It has been shown that the H44 type strain has two complete flagellin genes, being H4 and H44 [Ratiner, Y. A. (1998) "New flagellin specifying genes in some *E. coli* strains" J. Bacteriol 180: 979-984]. Thus, we conclude that all the three flagellin genes (obtained from type strains for H4, H17 and H44, and sequenced) encode the H4 flagellin, and that the flagellin genes for H17 and H44 specificities have not yet been cloned.

#### 10.5 *Flagellin gene coding for H10:*

The flagellin genes obtained from type strains for H10 and H50 are nearly identical, with 3 a.a. difference in the gene products. Strain P5560 carrying plasmid pPR1923 (which carries a flagellin gene from the H10 type strain) agglutinated with anti-H10 serum. We conclude that the sequence obtained from the H10 type strain encodes the H10 flagellin. It is not clear if the sequence obtained from the H50 type strain encodes H10 or H50 (see below section for H50).

#### 10.6 *Flagellin gene coding for H38:*

The flagellin genes obtained from type strains for H38 and H55 are nearly identical, with only 1 a.a. difference in the gene products. Strain M2126 carrying plasmid pPR1984 (carrying the flagellin gene from the type strain H38) agglutinated with anti-H38 serum, but not with

anti-H55 serum. It also has been shown that the type strain for H55 has two complete flagellin genes coding for H55 and H38 specificities [Ratiner, Y. A. (1998) "New flagellin specifying genes in some *E. coli* strains" J. Bacteriol 180: 979-984]. Thus, we conclude that both

#### 10.7 Summary:

Flagellin genes coding for 39 H antigens have been identified, being those for specificities H1, H2, H4, H5, H6, H7, H9, H10, H11, H12, H14, H15, H16, H18, H19, H20, H21, H23, H24, H26, H27, H28, H29, H30, H31, H32, H33, H34, H38, H39, H41, H42, H43, H45, H46, H49, H51, H52, and H56.

#### 11. Comparison and alignment of the flagellin genes:

Programs Pileup [Devereux, J., Haeberli, P. and Smithies, O.: A comprehensive set of sequence analysis programs for the VAX. Nucl. Acids Res. 12 (1984) 387-395] and Multicomp [Reeves, P.R., Farnell, L. and Lan, R.: MULTICOMP: a program for preparing sequence data for phylogenetic analysis. CABIOS 10 (1994) 281-284] were used.

The previously published sequence of H1 (GenBank accession number L07387) was extracted from GenBank and used. Because we did not sequence H36 and H53 flagellin genes and we did not have the H16 type strain, we only compared 51 flagellin genes of H type strains and the *fliC* genes from the additional 10 H7 strains.

Among the H7 *fliC* genes, the percentage of DNA difference ranged from 0.0 to 2.39%. The flagellin genes from type strains for H40 and H8 are identical. Some others are nearly identical: H21 and H47 (1.5% difference), H12 and H1 (2.6% difference), H10 and H50 (0.3% difference), H38 and H55 (0.1% difference), H4, H44 and H17 are very similar, the pairwise difference ranging from 0.33% to 0.87%.

For the flagellin genes obtained from type strains for H4, H17 and H44, we have shown that all the three genes encode flagellin with the H4 specificity (see above). For the flagellin genes obtained from type strains fro H21 and H47, and H38 and H55, we have confirmed the specificities for one for each pair and have good reason to conclude that both genes of each pair encode the same H specificity (see above section), being that for H21 and H38 specificities respectively.

For the flagellin genes obtained from type strains for H10 and H50, we have confirmed that the one from the H10 type strain encodes H10 specificity. As these two genes are highly similar, we have presumed that they both encode H10 specificity.

In the cases where the flagellin gene from two type strains is near identical, we conclude that both genes code for flagellin of the same H specificity and that one or other strain has an additional locus which carries the functional gene, although the flagellin genes sequenced do not appear to be mutated.

We have shown by cloning and expression that the flagellin genes obtained from the H1 and H12 type strains encode H1 and H12 specificities respectively (see above section). The neucleotide difference between these two genes is higher at 2.6% (see above), but still within the normal range for variation within a gene in *E. coli*. The two antigens cross react, and this cross reaction must be due to the high level similarity of the flagellins encoded by these two genes.

As discussed above, genes encoding some H antigens have been shown to be located at loci other than *fliC*. H3, H36, H47, H53 have been shown to be at a locus called *flkA*, H44 and H55 at *fllA*, and H54 at *flmA* [Ratiner Y A (1998) "New flagellin-specifying genes in some *Escherichia coli* strains" *J. Bacteriol.* 180 979-984]. However, these strains may carry a *fliC* in addition to *flkA*, *fllA* or *flmA* [Ratiner Y A (1998) "New flagellin-specifying genes in some

*Escherichia coli* strains" J. Bacteriol. 180 979-984].

The flagellin gene encoding H48 was previously sequenced from *E. coli* strain K-12 [Ku wajima G, Asaka J, Fujiwara T, Node K and Kondo E "Nucleotide sequence of the hag gene encoding flagellin of *Escherichia coli*" J Bacteriol. 168 (1986) 1479-1483]. We have sequenced the *fliC* gene from the H48 type strain, and found that it is identical to that from K-12.

The H54 gene is known to be at *flmA* [Ratiner Y A (1998) "New flagellin-specifying genes in some *Escherichia coli* strains" J. Bacteriol. 180 979-984]

and the finding of a non-functional presumptive *fliC* locus in the H54 strain shows that it is present but not expressed. However, we have not amplified and sequenced the functional *flmA* gene of this strain.

Using the 43 unique sequences (being the 39 identified genes with confirmed specificities and the flagellin genes obtained from the H8 (or H40), H25, H37, and H48 type strains) and the sequences from the two non-functional flagellin genes (from H type strains H35 and H54) (see Table 3) we have been able to determine antigen specific primers for each of the H antigen specificities and thereby show that it is practicable to detect *E.coli* strains carrying specific H antigens without false positives from strains of other H types. There is no reason to expect that the addition of 11 sequences to the 43 unique sequences obtained will affect the general conclusion, as unlike previous reports, our study covers flagellin sequences for a substantial majority of known *E. coli* H antigen specificities.

Our study of 11 H7 genes from strains of eight different O antigens shows limited variation which was such that the variation within genes for H antigens does not affect the ability to select antigen specific primers. O:H combinations in general define a strain and as some of the strains thus defined were quite distant from each other in a study by Whittam [Whittam T S, wolfe M L,

Wachsmuth I K, Orskov I and Wilson R A "Clonal relationships among *Escherichia coli* strains that cause hemorrhagic colitis and infantile diarrhea" Infect. Immun. 61 (1993) 1619-1629] the variation we observe is thought to represent that generally present in H7 genes. We also obtained more than one sequences for flagellin genes for H specificities H4, H10, and H38, and again the level of variation within a given specificities is very low. However, there is a low possibility that primers chosen without knowledge of the variation within genes of each H specificity could fail to give positive results with some isolates due to chance choice of primers which cover a base or bases which contribute to this low level variation. The variation within the H7 genes is in the normal range for variation within a gene in *E. coli* and if this possibility did occur it would be easy to use an alternate primer pair. For example, if a first primer in a primer pair is unable to hybridise to a target region because of low level variation in that region, a positive result may be achieved by using a second primer in that pair together with a third primer, whether or not the third primer is specific for the flagellin gene. Where the third primer is not specific for the flagellin gene, the specificity of the primer pair derives from the specificity of the second primer. The observation that the overall level of variation within gene for a given H specificity is very low making it extremely unlikely that the regions covered by the two primers specific for H specificity would both have undergone change in the same strain.

There are 54 known H antigens for *E. coli* and of these there are 11 H antigen specificities for which we do not as yet have sequence. It will be easy to determine these sequences and determine primer pairs specific for these H antigens by comparing these sequences with the 45 obtained sequences (see Table 3), and also modify the primers selected for any H antigen for which we already

know the sequence in the unlikely event that there is a possibility of false positives with the primers selected.

The sequences for the remaining H antigens can be obtained in one of the following ways:

5 1. where we have two bands by PCR (H36 and H53 type strains), we purify each and sequence, and also clone each into a strain mutated in its *fliC* gene and determine the H antigen expressed by use of specific sera. In this way a  
10 specific sequence can be related to an H antigen specificity. The other band which represents an H antigen gene for a different specificity is expected to include a mutant gene or a gene similar to one of those for a known H specificity, but if not may represent a new specificity for which primer pairs could be selected. It may be difficult to obtain expression of flagellin genes when  
15 cloned from *E. coli* due to cloning together with regulatory sequences which prevent expression. This is easily avoided by cloning the major segment of the gene into a functioning *fliC* gene to replace the equivalent segment of that gene, using standard site directed mutagenesis to give suitable restriction sites within the  
20 cloned gene and incorporating those restriction sites into primers used to amplify the major segment of the gene to be studied to facilitate the cloning. We have cloned and  
25 sequenced the PCR bands from the H36 and the H55 type strains using this method (see section 16).

30 2. Where two or more strains have the same flagellin gene sequence, the genes are cloned as above and the H antigen specificity represented by this sequence is determined. This identifies the strain in which the expected gene is expressed and also those strains for which we have sequenced a gene which is not being  
35 expressed. We then clone the gene for the antigen expressed in these strains by making a bank of plasmid clones using chromosomal DNA and select for a clone which

is expressing an H antigen different from the one represented by the known sequence. This can be done by taking advantage of the fact that the H antigen is on flagellin, the protein of the bacterial flagellum used for movement of the bacteria. In the presence of antibodies specific to that flagellum the bacteria cannot swim. For selection the clones are placed in a situation in which motile cells can swim away from the others and be collected. There are many versions of these techniques and any could be used. One version is to place the bacteria on a nutrient agar plate with reduced agar content such that bacteria can swim away from the site of inoculation. This is easily seen as growth on the plate and a sample of the bacteria which are motile can be recovered and cultivated. In this way bacteria carrying cloned H antigen genes can be selected. If the medium in the plate has antibody added to it only bacteria which express an H antigen different to that recognised by the antiserum will be able to swim. Specifically if the antiserum used is specific for the H antigen expressed by the gene for which we have sequence, only clones which express a different H antigen, such as those expressing the H antigen expressed by the H type strains used to make the plasmid, will be selected. Once the clone is obtained, the H antigen gene can be sequenced.

Our work has shown that there are at least 7 cases where the H antigen type strains carry two H antigen genes which appear to be complete and have the potential to function. However, while *E. coli* does not (in general) have a capacity to express more than one flagellin gene, it is striking that there are several loci for flagellin genes [Ratiner Y A (1998) "New flagellin-specifying genes in some *Escherichia coli* strains" J. Bacteriol. 180 979-984]. Several of the pairs of H type strains with identical or near identical sequence do not include any of the H antigen types shown by Ratiner [Ratiner Y A

(1998) "New flagellin-specifying genes in some *Escherichia coli* strains" J. Bacteriol. 180 979-984] to map other than at *fliC* although these predominate. This suggests that there are additional cases where the expressed gene is not the only flagellin gene present. However the fact that many of the cases where we obtained flagellin genes of identical or near identical sequence and/or two flagellin genes from one strain involve type strains found by Ratiner [Ratiner Y A (1998) "New flagellin-specifying genes in some *Escherichia coli* strains" J. Bacteriol. 180 979-984] to map away from *fliC* are among those near identical to others, indicates that the phenomenon is of limited extent. Nonetheless it remains possible even where only one gene has been obtained by PCR, that it is one of a pair of flagellin genes, the other not being amplified by the primers used, and further that it is the one not amplified which is expressing the H antigen of the strain.

It will therefore be necessary to clone as described above each of the flagellin genes we have sequenced and confirm that it expresses the expected antigen to ensure that the invention give results corresponding to those of the traditional serotyping scheme. In the event that it does not, the gene for the type antigen can be cloned and sequenced by the means described above.

The 11 H7 *fliC* sequences fell into three groups, one comprising the genes from the O157:H7 and O55:H7 strains, which were identical, as expected given the proposed relationship between the clones. It has been shown that *E. coli* O157:H7 and O55:H7 clones are closely related [Whittam T S, wolfe M L, Wachsmuth I K, Orskov I and Wilson R A "Clonal relationships among *Escherichia coli* strains that cause hemorrhagic colitis and infantile diarrhea" Infect. Immun. 61 (1993) 1619-1629] thus it was expected that the H7 *fliC* genes from O157 and O55 would be identical. Among the H7 *fliC* sequences, we can identify primers specific to the H7 *fliC* gene for each of the three H7 groups. Two of these primers in combination with an H7



specific primer gave two primer pairs specific for the H7 gene of from the O157:H7 and O55:H7 clones.

13. *Specific oligonucleotide primers for each of the 43 flagellin genes*

Two oligonucleotide primers were chosen based on each of the 43 sequences. None of them had more than 85% identity with any other of 61 flagellin gene sequences. Thus, these primers are specific for each H type. These primers are listed in Table 3.

The flagellin gene of the H54 type strain is a mutated gene. It has an insertion sequence (IS1222) inserted into a normal flagellin gene of H21. Thus, primers for H21 would amplify a fragment of different size in H54. We also provide 2 primers based on the insertion sequence (see H54 row in Table 3), and the use of one of them in combination with one of the H21 primers will generate a PCR band only in H54, which will also differentiate those strain carrying the mutated H21 gene from those expressing the H21 flagellin gene.

The *flic* gene of H35 type strain is also a mutated gene. It has an insertion sequence (IS1) inserted into a normal flagellin gene of H11. Thus, primers for H11 would amplify a fragment of different size in H35. We also provide 2 primers based on the insertion sequence (see H35 row in Table 3), and the use of one of them in combination with one of the H11 primers will generate a PCR band only in H35, which will also differentiate those strain carrying the mutated H11 gene from those expressing the H11 flagellin gene.

14. *Testing of the H7 specific oligonucleotide primers*

Primer pair #1806/#1809 (see Table 3) was used to carry out PCR on chromosomal DNA samples of all the 54 H type strains and the H7 strains listed in Table 1. PCR reactions were carried out under the following conditions: denaturing, 94°C/30'; annealing, 58°C/30'; extension,

72°C/1'; 30 cycles. PCR reaction was carried out in an volume of 50ul for each of the chromosomal sample. After the PCR reaction, 5ul PCR product from each sample was run on an agarose gel to check for amplified DNA.

5 Primer pairs #1806/#1809 produced a band of predicted size with all the 11 strains expressing H7, but gave no band with other H type strains. Thus, these primers are H7 specific.

10 15. *Testing of oligonucleotide primers specific to H7 of O157 and O55:*

Based on a comparison of the *fliC* sequences of 11 different H7 strains, we have identified two oligonucleotides [#1696 (5'-GGCCTGACTCAGGCGGCC) at positions 178 to 195 in M527 and #1697 (5'-GAGTTACCGCGCTGCTGA) positions 1700-1683 in M527] which are unique to H7 of O157 and O55. Although not identical to any parts of the *fliC* sequences of any other H7 strains, these two primers are identical or have high level similarity to *fliC* genes of some other H types. However a combination of one of these primers with one of the H7 specific primers can give specificity for H7 of O157:H7 and O55:H7 *E. coli*.

25 Primer pairs #1696/#1809 and #1697/#1806 were used to carry out PCR on chromosomal DNA samples of all the H type strains and the H7 strains listed in Table 1. PCR reactions were carried out under the following conditions: denaturing, 94°C/30'; annealing, 61°C/30' (for #1696/#1809) or 60°C/30' (for #1697/#1806); extension, 72°C/1'; 30 cycles. PCR reaction was carried out in an volume of 50ul for each of the chromosomal samples. After the PCR reaction, 5ul PCR product from each sample was run on an agarose gel to check for amplified DNA.

35 Both primer pairs produced a band of predicted size with both of the O157:H7 strains (strains M1004 and M527, see Table 1), and the O55:H7 strain (strain M1686, see Table 1), but gave no band with other strains. Thus, these

- 44 -

two pairs of primers are specific to H7 genes of O157:H7 and O55:H7 *E. coli* strains.

16. Identification of flagellin genes for the remaining 15 *H* specificities.

16.1. Sequencing the potential *flkA* gene coding for the H36 flagellin:

Using primers #1431 (5'- atg gca caa gtc att aat acc caa c) and #1432 (5'- cta acc ctg cag cag aga ca), we have amplified two bands from the H36 type strain. PCR reaction was carried out under the following conditions: denaturing, 94°C/30'; annealing, 57°C/30'; extension, 72°C/1'; 30 cycles. These two PCR fragments were then cloned into the pGEM-T vector using the Promega pGEM-T cloning kit (Madison WI USA) to make plasmids pPR1992 and pPR1993. Inserts from both plasmids were first sequenced using the M13 universal primers (which bind to the pGEM-T DNA flanking the insertion site). For pPR1992, primers based on the sequence obtained were then used to sequence further, and this procedure was repeated until the insert was fully sequenced.

The sequence of the insert of pPR1992 is identical to that of the H12 flagellin gene sequence except perhaps for the first 8 and last 7 codons which are encoded by the PCR primers in plasmid pPR1992. We have only sequenced the two ends of the insert of plasmid pPR1993 (Figures 71 and 72), and the sequences of the two ends of the insert of pPR1993 are very similar to ends of other sequenced flagellin genes. We conclude that the insert of plasmid pPR1993 encodes a flagellin gene. The full sequence of the insert of plasmid pPR1993 can be obtained using the same method as for the sequencing of the insert of plasmid pPR1992. It is known that *flkA* gene encodes the H36 flagellin [Ratiner, Y. A. (1998) "New flagellin specifying genes in some *E. coli* strains" *J. Bacteriol* 180: 979-984], and it is highly likely that plasmid pPR1993 contains the

*flkA* gene of the H36 type strain. H specificities can be confirmed by slide agglutination.

The currently uncharacterised sequence of both ends and of DNA flanking these two sequenced genes can be obtained by PCR walking and sequencing. Methods for PCR walking from a known sequence to an unknown region in chromosomal DNA are available (see [Siebert, P. D. , A. Chenchi, D. E. Kellogg, A. Lukyanov and S. A. Lukyanov (1995) "An improved PCR method for walking in uncloned genomic DNA." Nuc. Acids Res. 23: 1087-1088]).

The sequenced genes then can be PCR amplified and cloned using the method(s) described in section 9. Flagellins expressed by strain M2126 carrying these plasmids then can be determined by use of specific sera.

The sequences flanking the *flkA* gene can then be used to PCR amplify other *flkA* genes (see below).

#### 16.2 The *flkA* genes coding for H3, H47 and H53:

It has been shown that flagellins H3, H47 and H53 are encoded by *flkA* genes in the type strains [Ratiner, Y. A. (1998) "New flagellin specifying genes in some *E. coli* strains" J. Bacteriol 180: 979-984]. These genes can be PCR amplified using primers based on the sequences flanking the *flkA* gene in the H36 type strain. These PCR fragments can then be sequenced, and the genes expressed in strain M2126 for the identification of these genes.

#### 16.3 The *flfA* genes coding for H44 and H55:

It is known that flagellins H44 and H55 are coded by *flfA* genes.

##### 16.3.1 The H55 flagellin gene:

Using primers #1868 and #1870 (Table 3B), we have amplified two bands from the H55 type strain. PCR reaction was carried out under the following conditions: denaturing, 94°C/30'; annealing, 50°C/30'; extension, 72°C/1'; 30 cycles. These two PCR fragments were then

cloned into the pGEM-T vector using the Promega pGEM-T cloning kit (Madison WI USA) to make plasmids pPR1994 and pPR1989. Inserts from both plasmids were first sequenced using the M13 universal primers (which bind to the pGEM-T DNA flanking the insertion site). Primers based on the sequence obtained were then used to sequence further, and this procedure was repeated until both inserts were fully or partly sequenced.

The sequence of the insert of pPR1994 is highly similar to that of the flagellin gene of the H38 type strain, with 1 amino acid difference in the gene products. We have only sequenced the two ends of the insert of plasmid pPR1989 (figures 70A and 70B), and the sequences of the two ends of the insert of pPR1989 are very similar to ends of other sequenced flagellin genes. We conclude that the insert of plasmid pPR1989 encodes a flagellin gene. The full sequence of the insert of plasmid pPR1989 can be obtained using the same method as for the sequencing of the insert of plasmid pPR1994. It is known that the H55 type strain carries flagellin genes for both H38 and H55, and that the H55 flagellin gene is at the *flaA* locus [Ratiner, Y. A. (1998) "New flagellin specifying genes in some *E. coli* strains" J. Bacteriol 180: 979-984]. Thus, it is highly likely that plasmid pPR1989 contains the *flaA* gene of the H55 type strain.

The currently uncharacterised sequence of both ends and of DNA flanking these two sequenced genes can be obtained by PCR walking and sequencing. Methods for PCR walking from a known sequence to an unknown region in chromosomal DNA are available (see [Siebert, P. D. , A. Chenchi, D. E. Kellogg, A. Lukyanov and S. A. Lukyanov (1995) "An improved PCR method for walking in uncloned genomic DNA." Nuc. Acids Res. 23: 1087-1088]).

The sequenced genes then can be PCR amplified and cloned using the method(s) described in section 9. Flagellins expressed by strain M2126 carrying these plasmids then can be determined by use of specific sera.

### 16.3.2 The H44 flagellin gene:

The sequence information for DNA flanking the *flIA* gene in the H55 type strain can then be used to PCR, sequence and identify the *flIA* gene in the H44 type strain.

### 16.4 The *flmA* gene coding for H54:

This gene can be cloned by making a bank of plasmid clones in strain M2126 using chromosomal DNA of the H54 type strain and selecting for a transformant which is motile on an agar plate. This is done by taking advantage of the fact that the H antigen is on flagellin, the protein of the bacterial flagellum used for movement of the bacteria. Strain M2126 lacks flagellin. Once the clone(s) is obtained and identified by use of anti-H54 serum, the flagellin gene can be sequenced. It is possible that clones expressing different flagellin specificities can be obtained, and each of them can be identified by using different sera.

### 16.5 The flagellin genes obtained from the H37 and H48 type strains:

We have used primers #1868 and #1869 (both were based on the sequence obtained from the H48 type strain, also see section 9) and primers #1868 and #1870 (both were based on the sequences of the H7 flagellin gene of the H7 type strain, also see section 9) to PCR amplify and clone the sequenced flagellin genes from the H48 and H37 type strains respectively. Strain P5560 carrying the plasmid containing either the cloned gene was not motile and did not react with the appropriate antisera. It is highly likely that mutations have occurred due to PCR errors. This can be resolved by re-amplification and re-cloning of the genes.

### 16.6 The flagellin gene obtained from the H25 type

*strain:*

The flagellin gene sequence we first obtained from the H25 type strain lacks 23 and 21 codons at 5' and 3' ends respectively. We could not amplify the full gene from the H25 type strain using primers based on the H7 flagellin gene of the H7 type strain, and it was necessary to get the full sequence of this flagellin gene by other means.

We have used primers (#2650: 5' - cag cga tga aat act tgc cat and #2648: 5' - caa tgc ttc gtg acg cac) based on the genes (*fliD* and *fliA* respectively) flanking *fliC* gene in *E. coli* K-12 [Blattner, P. R., G. I. Plunkett, C. A. Bloch, N. T. Perna, V. Burland, M. Riley and et al. (1997) "The complete genome sequence of *E. Coli* K12" Science 277: 1453-1474] and primers (#2658: 5' - gcc tga gtc aga cct ttg and # 2653 5' - aac ctg tct gaa gcg cag) based on the flagellin sequence obtained from the H25 type strain to PCR amplify both ends of the flagellin gene. The PCR product was then sequenced, and we have now obtained the full flagellin gene sequence and sequence for the DNA flanking the flagellin gene from type strain H25 (Figure 69). Now, it is straightforward to PCR amplify, clone and express, and identify this gene using the methods described in sections 9 and 10.

16.7      *The flagellin genes obtained from the H8 and H40 type strains:*

The flagellin gene sequences obtained from both the H8 and H40 type strains lack 18 and 15 codons at 5' and 3' ends respectively. We have used primers based on the H7 flagellin gene of the H7 type strain to PCR amplify and clone the full genes from these two strains. Strain M2126 carrying plasmid made this way was not motile under microscope and did not react with the appropriate antisera. This could be due to PCR errors as mentioned in section 16.5 or perhaps the first and last few amino acids encoded by the primers (based on H7 flagellin gene) are

incompatible in this case.

The full sequence of the full gene can be obtained using method described in section 16.6. The flagellin gene can then be PCR amplified, cloned and expressed, and identified using the methods described in sections 9 and 10.

The gene products of the flagellin genes obtained from the H8 and H40 type strains are identical. Thus, one of these two H specificities must be encoded by a unknown gene, and it can be cloned and identified using the method described in the section 16.8.

#### 16.8 *Flagellin genes coding for H17, H35, and H50:*

As mentioned above, the sequenced flagellin genes from the H17 and H50 type strains encode H4 and H10 specificities respectively. The flagellin gene sequence obtained from the H35 strain has a insertion and encodes a non-functional gene (see section 8). Thus, genes coding for these flagellins have not been identified, and their location is unknown. One can use primers based on DNA flanking *fliC*, *fliA*, *flkA*, and *flmA* to do PCR on the type strain for each of the flagellin antigen. PCR products can then be sequenced, and possible genes can be cloned, expressed and identified then.

If the target gene is not PCR amplified using primers based on sequence of these loci or sequence flanking these loci, it can be cloned by making a bank of plasmid clones in strain M2126 using chromosomal DNA of the type strain and selecting for a transformant which is motile on an agar plate. This is done by taking advantage of the fact that the H antigen is on flagellin, the protein of the bacterial flagellum used for movement of the bacteria. Strain M2126 lacks flagellin. Once the clone(s) is obtained and identified by use of antisera, the flagellin gene can be sequenced. It is possible that clones expressing different flagellin antigens can be obtained,



and each of them can be identified by using different antisera. Antiserum for H50 can be prepared using standard methods [Ewing, W.H.: Edwards and Ewing's identification of the *Enterobacteriaceae*., Elsevier Science Publishers, Amsterdam, The Netherlands, 1986].

## O antigen

### Materials and Methods-part 1

The experimental procedures for the isolation and characterisation of the *E. coli* O111 O antigen gene cluster (position 3,021-9,981) are according to Bastin D.A., et al. 1991 "Molecular cloning and expression in *Escherichia coli* K-12 of the *rfb* gene cluster determining the O antigen of an *E. coli* O111 strain". *Mol. Microbiol.* 5:9 2223-2231 and Bastin D.A. and Reeves, P.R. 1995 "Sequence and analysis of the O antigen gene(*rfb*) cluster of *Escherichia coli* O111". *Gene* 164: 17-23.

#### A. Bacterial strains and growth media

Bacteria were grown in Luria broth supplemented as required.

#### B. Cosmids and phage

Cosmids in the host strain x2819 were repackaged *in vivo*. Cells were grown in 250mL flasks containing 30mL of culture, with moderate shaking at 30°C to an optical density of 0.3 at 580 nm. The defective lambda prophage was induced by heating in a water bath at 45°C for 15min followed by an incubation at 37°C with vigorous shaking for 2hr. Cells were then lysed by the addition of 0.3mL chloroform and shaking for a further 10min. Cell debris were removed from 1mL of lysate by a 5min spin in a microcentrifuge, and the supernatant removed to a fresh microfuge tube. One drop of chloroform was added then shaken vigorously through the tube contents.

#### C. DNA preparation

Chromosomal DNA was prepared from bacteria grown overnight at 37°C in a volume of 30mL of Luria broth. After harvesting by centrifugation, cells were washed and

resuspended in 10mL of 50mMTris-HCl pH 8.0. EDTA was added and the mixture incubated for 20min. Then lysozyme was added and incubation continued for a further 10min. Proteinase K, SDS, and ribonuclease were then added and the mixture incubated for up to 2hr for lysis to occur. All incubations were at 37°C. The mixture was then heated to 65°C and extracted once with 8mL of phenol at the same temperature. The mixture was extracted once with 5mL of phenol/chloroform/iso-amyl alcohol at 4°C. Residual phenol was removed by two ether extractions. DNA was precipitated with 2 vols. of ethanol at 4°C, spooled and washed in 70% ethanol, resuspended in 1-2mL of TE and dialysed. Plasmid and cosmid DNA was prepared by a modification of the Birnboim and Doly method [Birnboim, H. C. and Doly, J. (1979) "A rapid alkaline extraction procedure for screening recombinant plasmid DNA" *Nucl. Acid Res.* 7:1513-1523]. The volume of culture was 10mL and the lysate was extracted with phenol/chloroform/iso-amyl alcohol before precipitation with isopropanol. Plasmid DNA to be used as vector was isolated on a continuous caesium chloride gradient following alkaline lysis of cells grown in 1L of culture.

D. Enzymes and buffers.

Restriction endonucleases and DNA T4 ligase were purchased from Boehringer Mannheim (Castle Hill, NSW, Australia) or Pharmacia LKB (Melbourne, VIC Australia). Restriction enzymes were used in the recommended commercial buffer.

E. Construction of a gene bank.

Individual aliquots of M92 chromosomal DNA (strain Stoke W, from Statens Serum Institut, 5 Artillerivej, 2300 Copenhagen S, Denmark) were partially digested with 0.2U *Sau3A1* for 1-15mins. Aliquots giving the greatest proportion of fragments in the size range of approximately 40-50kb were selected and ligated to vector pPR691 previously digested with *Bam*H1 and *Pvu*II. Ligation mixtures were packaged *in vitro* with packaging extract.

The host strain for transduction was x2819 and recombinants were selected with kanamycin.

#### F. Serological procedures.

Colonies were screened for the presence of the O111 antigen by immunoblotting. Colonies were grown overnight, up to 100 per plate then transferred to nitrocellulose discs and lysed with 0.5N HCl. Tween 20 was added to TBS at 0.05% final concentration for blocking, incubating and washing steps. Primary antibody was *E. coli* O group 111 antiserum, diluted 1:800. The secondary antibody was goat anti-rabbit IgG labelled with horseradish peroxidase diluted 1:5000. The staining substrate was 4-chloro-1-naphthol. Slide agglutination was performed according to the standard procedure.

#### G. Recombinant DNA methods.

Restriction mapping was based on a combination of standard methods including single and double digests and sub-cloning. Deletion derivatives of entire cosmids were produced as follows: aliquots of 1.8mg of cosmid DNA were digested in a volume of 20ml with 0.25U of restriction enzyme for 5-80min. One half of each aliquot was used to check the degree of digestion on an agarose gel. The sample which appeared to give a representative range of fragments was ligated at 4°C overnight and transformed by the  $\text{CaCl}_2$  method into JM109. Selected plasmids were transformed into sf174 by the same method. P4657 was transformed with pPR1244 by electroporation.

#### H. DNA hybridisation

Probe DNA was extracted from agarose gels by electroelution and was nick-translated using [ $\alpha$ - $^{32}\text{P}$ ]-dCTP. Chromosomal or plasmid DNA was electrophoresed in 0.8% agarose and transferred to a nitrocellulose membrane. The hybridisation and pre-hybridisation buffers contained either 30% or 50% formamide for low and high stringency probing respectively. Incubation temperatures were 42°C and 37°C for pre-hybridisation and hybridisation respectively. Low stringency washing of

filters consisted of 3 x 20min washes in 2 x SSC and 0.1% SDS. High-stringency washing consisted of 3 x 5min washes in 2 x SSC and 0.1% SDS at room temperature, a 1hr wash in 1 x SSC and 0.1% SDS at 58°C and 15min wash in 0.1 x SSC and 0.1% SDS at 58°C.

I. Nucleotide sequencing of *E. coli* O111 O antigen gene cluster (position 3,021-9,981)

Nucleotide sequencing was performed using an ABI 373 automated sequencer (CA, USA). The region between map positions 3.30 and 7.90 was sequenced using uni-directional exonuclease III digestion of deletion families made in PT7T3190 from clones pPR1270 and pPR1272. Gaps were filled largely by cloning of selected fragments into M13mp18 or M13mp19. The region from map positions 7.90-10.2 was sequenced from restriction fragments in M13mp18 or M13mp19. Remaining gaps in both the regions were filled by priming from synthetic oligonucleotides complementary to determined positions along the sequence, using a single stranded DNA template in M13 or phagemid. The oligonucleotides were designed after analysing the adjacent sequence. All sequencing was performed by the chain termination method. Sequences were aligned using SAP [Staden, R., 1982 "Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing". *Nuc. Acid Res.* 10: 4731-4751; Staden, R., 1986 "The current status and portability of our sequence handling software". *Nuc. Acid Res.* 14: 217-231]. The program NIP [Staden, R. 1982 "An interactive graphics program for comparing and aligning nucleic acid and amino acid sequence". *Nuc. Acid Res.* 10: 2951-2961] was used to find open reading frames and translate them into proteins.

J. Isolation of clones carrying *E. coli* O111 O antigen gene cluster

The *E. coli* O antigen gene cluster was isolated according to the method of Bastin D.A., et al. [1991 "Molecular cloning and expression in *Escherichia coli* K-

12 of the *rfb* gene cluster determining the O antigen of an *E. coli* O111 strain". *Mol. Microbiol.* 5(9), 2223-2231]. Cosmid gene banks of M92 chromosomal DNA were established in the *in vivo* packaging strain x2819. From the genomic bank,  $3.3 \times 10^3$  colonies were screened with *E. coli* O111 antiserum using an immuno-blotting procedure: 5 colonies (pPR1054, pPR1055, pPR1056, pPR1058 and pPR1287) were positive. The cosmids from these strains were packaged *in vivo* into lambda particles and transduced into the *E. coli* deletion mutant Sfl74 which lacks all O antigen genes. In this host strain, all plasmids gave positive agglutination with O111 antiserum.

An *Eco* RI restriction map of the 5 independent cosmids showed that they have a region of approximately 11.5 kb in common (Figure 1). Cosmid pPR1058 included sufficient flanking DNA to identify several chromosomal markers linked to O antigen gene cluster and was selected for analysis of the O antigen gene cluster region.

#### K. Restriction mapping of cosmid pPR1058

Cosmid pPR1058 was mapped in two stages. A preliminary map was constructed first, and then the region between map positions 0.00 and 23.10 was mapped in detail, since it was shown to be sufficient for O111 antigen expression. Restriction sites for both stages are shown in Figure 2. The region common to the five cosmid clones was between map positions 1.35 and 12.95 of pPR1058.

To locate the O antigen gene cluster within pPR1058, pPR1058 cosmid was probed with DNA probes covering O antigen gene cluster flanking regions from *S. enterica* LT2 and *E. coli* K-12. Capsular polysaccharide (*cps*) genes lie upstream of O antigen gene cluster while the gluconate dehydrogenase (*gnd*) gene and the histidine (*his*) operon are downstream, the latter being further from the O antigen gene cluster. The probes used were pPR472 (3.35kb), carrying the *gnd* gene of LT2, pPR685 (5.3kb) carrying two genes of the *cps* cluster, *cpsB* and

*cpsG* of LT2, and K350 (16.5kb) carrying all of the *his* operon of K-12. Probes hybridised as follows: pPR472 hybridised to 1.55kb and 3.5 kb (including 2.7 kb of vector) fragments of *Pst*I and *Hind*III double digests of pPR1246 (a *Hind*III/*Eco*R1 subclone derived from pPR1058, Figure 2), which could be located at map positions 12.95-15.1; pPR685 hybridised to a 4.4 kb *Eco*R1 fragment of pPR1058 (including 1.3 kb of vector) located at map position 0.00-3.05; and K350 hybridised with a 32kb *Eco*R1 fragment of pPR1058 (including 4.0kb of vector), located at map position 17.30-45.90. Subclones containing the presumed *gnd* region complemented a *gnd<sup>-</sup>edd<sup>-</sup>* strain GB23152. On gluconate bromothymol blue plates, pPR1244 and pPR1292 in this host strain gave the green colonies expected of a *gnd<sup>-</sup>edd<sup>-</sup>* genotype. The *his<sup>+</sup>* phenotype was restored by plasmid pPR1058 in the *his* deletion strain Sf174 on minimal medium plates, showing that the plasmid carries the entire *his* operon.

It is likely that the O antigen gene cluster region lies between *gnd* and *cps*, as in other *E. coli* and *S. enterica* strains, and hence between the approximate map positions 3.05 and 12.95. To confirm this, deletion derivatives of pPR1058 were made as follows: first, pPR1058 was partially digested with *Hind*III and self ligated. Transformants were selected for kanamycin resistance and screened for expression of O111 antigen. Two colonies gave a positive reaction. *Eco*R1 digestion showed that the two colonies hosted identical plasmids, one of which was designated pPR1230, with an insert which extended from map positions 0.00 to 23.10. Second pPR1058 was digested with *Sal*I and partially digested with *Xho*I and the compatible ends were re-ligated. Transformants were selected with kanamycin and screened for O111 antigen expression. Plasmid DNA of 8 positively reacting clones was checked using *Eco*R1 and *Xho*I digestion and appeared to be identical. The cosmid of one was designated pPR1231. The insert of pPR1231

contained the DNA region between map positions 0.00 and 15.10. Third, pPR1231 was partially digested with *Xho*I, self-ligated, and transformants selected on spectinomycin/ streptomycin plates. Clones were screened for kanamycin sensitivity and of 10 selected, all had the DNA region from the *Xho*I site in the vector to the *Xho*I site at position 4.00 deleted. These clones did not express the O111 antigen, showing that the *Xho*I site at position 4.00 is within the O antigen gene cluster. One clone was selected and named pPR1288. Plasmids pPR1230, pPR1231, and pPR1288 are shown in Figure 2.

L. Analysis of the *E. coli* O111 O antigen gene cluster (position 3,021-9,981) nucleotide sequence data

Bastin and Reeves [1995 "Sequence and analysis of the O antigen gene(*rfb*)cluster of *Escherichia coli* O111". Gene 164: 17-23] partially characterised the *E.coli* O111 O antigen gene cluster by sequencing a fragment from map position 3,021-9,981. Figure 3 shows the gene organisation of position 3,021-9,981 of *E. coli* O111 O antigen gene cluster. *orf3* and *orf6* have high level amino acid identity with *wcaH* and *wcaG* (46.3% and 37.2% respectively), and are likely to be similar in function to sugar biosynthetic pathway genes in the *E. coli* K-12 colanic gene cluster. *orf4* and *orf5* show high levels of amino acid homology to *manC* and *manB* genes respectively. *orf7* shows high level homology with *rfbH* which is an abequeose pathway gene. *orf8* encodes a protein with 12 transmembrane segments and has similarity in secondary structure to other *wzx* genes and is likely therefore to be the O antigen flippase gene.

Materials and Methods-part 2

A. Nucleotide sequencing of 1 to 3,020 and 9,982 to 14,516 of the *E. coli* O111 O antigen gene cluster

The sub clones which contained novel nucleotide sequences, pPR1231 (map position 0 and 1,510), pPR1237 (map position -300 to 2,744), pPR1239 (map position 2,744

to 4,168), pPR1245 (map position 9,736 to 12,007) and pPR1246 (map position 12,007 to 15,300) (Figure 2), were characterised as follows: the distal ends of the inserts of pPR1237, pPR1239 and pPR1245 were sequenced using the M13 forward and reverse primers located in the vector. PCR walking was carried out to sequence further into each insert using primers based on the sequence data and the primers were tagged with M13 forward or reverse primer sequences for sequencing. This PCR walking procedure was repeated until the entire insert was sequenced. pPR1246 was characterised from position 12,007 to 14,516. The DNA of these sub clones was sequenced in both directions.

The sequencing reactions were performed using the dideoxy termination method and thermocycling and reaction products were analysed using fluorescent dye and an ABI automated sequencer (CA, USA).

B. Analysis of the *E. coli* O111 O antigen gene cluster (positions 1 to 3,020 and 9,982 to 14,516 of Figure 5) nucleotide sequence data

The gene organisation of regions of *E. coli* O111 O antigen gene cluster which were not characterised by Bastin and Reeves [1995 "Sequence and analysis of the O antigen gene(*rfb*)cluster of *Escherichia coli* O111." *Gene* 164: 17-23], (positions 1 to 3,020 and 9,982 to 14,516) is shown in Figure 3. There are two open reading frames in region 1. Four open reading frames are predicted in region 2. The position of each gene is listed in Table 9.

The deduced amino acid sequence of *orf1* (*wbdH*) shares about 64% similarity with that of the *rfp* gene of *Shigella dysenteriae*. *Rfp* and *WbdH* have very similar hydrophobicity plots and both have a very convincing predicted transmembrane segment in a corresponding position. *rfp* is a galactosyl transferase involved in the synthesis of LPS core, thus *wbdH* is likely to be a galactosyl transferase gene. *orf2* has 85.7% identity at amino acid level to the *gmd* gene identified in the *E.*



*coli* K-12 colanic acid gene cluster and is likely to be a *gmd* gene. *orf9* encodes a protein with 10 predicted transmembrane segments and a large cytoplasmic loop. This inner membrane topology is a characteristic feature of all known O antigen polymerases thus it is likely that *orf9* encodes an O antigen polymerase gene, *wzy*. *orf10* (*wbdL*) has a deduced amino acid sequence with low homology with *Lsi2* of *Neisseria gonorrhoeae*. *Lsi2* is responsible for adding GlcNAc to galactose in the synthesis of lipooligosaccharide. Thus it is likely that *wbdL* is either a colitose or glucose transferase gene. *orf11* (*wbdM*) shares high level nucleotide and amino acid similarity with TrsE of *Yersinia enterocolitica*. TrsE is a putative sugar transferase thus it is likely that *wbdM* encodes the colitose or glucose transferase.

In summary three putative transferase genes and an O antigen polymerase gene were identified at map position 1 to 3,020 and 9,982 to 14,516 of *E. coli* O111 O antigen gene cluster. A search of GenBank has shown that there are no genes with significant similarity at the nucleotide sequence level for two of the three putative transferase genes or the polymerase gene. Figure 5 provides the nucleotide sequence of the O111 antigen gene cluster.

### Materials and Methods-part 3

A. PCR amplification of O157 antigen gene cluster from an *E. coli* O157:H7 strain (Strain C664-1992, from Statens Serum Institut, 5 Artillerivej, 2300, Copenhagen S, Denmark)

*E. coli* O157 O antigen gene cluster was amplified by using long PCR [Cheng et al. 1994, "Effective amplification of long targets from cloned inserts and human and genomic DNA" P.N.A.S. USA 91: 5695-5699] with one primer (primer #412: att ggt agc tgt aag cca agg gcg gta gcg t) based on the JumpStart sequence usually found in the promoter region of O antigen gene clusters [Hobbs,

et al. 1994 "The JumpStart sequence: a 39 bp element common to several polysaccharide gene clusters" Mol. Microbiol. 12: 855-856], and another primer #482 (cac tgc cat acc gac gac gcc gat ctg ttg ctt gg) based on the *gnd* gene usually found downstream of the O antigen gene cluster. Long PCR was carried out using the Expand Long Template PCR System from Boehringer Mannheim (Castle Hill NSW Australia), and products, 14 kb in length, from several reactions were combined and purified using the Promega Wizard PCR preps DNA purification System (Madison WI USA). The PCR product was then extracted with phenol and twice with ether, precipitated with 70% ethanol, and resuspended in 40mL of water.

B. Construction of a random DNase I bank:

Two aliquots containing about 150ng of DNA each were subjected to DNase I digestion using the Novagen DNase I Shotgun Cleavage (Madison WI USA) with a modified protocol as described. Each aliquot was diluted into 45ml of 0.05M Tris -HCl (pH7.5), 0.05mg/mL BSA and 10mM MnCl<sub>2</sub>. 5mL of 1:3000 or 1:4500 dilution of DNaseI (Novagen) (Madison WI USA) in the same buffer was added into each tube respectively and 10ml of stop buffer (100mM EDTA), 30% glycerol, 0.5% Orange G, 0.075% xylene and cyanol (Novagen) (Madison WI USA) was added after incubation at 15°C for 5 min. The DNA from the two DNaseI reaction tubes were then combined and fractionated on a 0.8% LMT agarose gel, and the gel segment with DNA of about 1kb in size (about 1.5mL agarose) was excised. DNA was extracted from agarose using Promega Wizard PCR Preps DNA Purification (Madison WI USA) and resuspended in 200 mL water, before being extracted with phenol and twice with ether, and precipitated. The DNA was then resuspended in 17.25 mL water and subjected to T4 DNA polymerase repair and single dA tailing using the Novagen Single dA Tailing Kit (Madison WI USA). The reaction product (85ml containing about 8ng DNA) was then extracted with chloroform:isoamyl alcohol (24:1) once and

ligated to  $3 \times 10^{-3}$  pmol pGEM-T (Promega) (Madison WI USA) in a total volume of 100mL. Ligation was carried out overnight at 4°C and the ligated DNA was precipitated and resuspended in 20mL water before being electroporated into *E. coli* strain JM109 and plated out on BCIG-IPTG plates to give a bank.

#### C. Sequencing

DNA templates from clones of the bank were prepared for sequencing using the 96-well format plasmid DNA miniprep kit from Advanced Genetic Technologies Corp (Gaithersburg MD USA). The inserts of these clones were sequenced from one or both ends using the standard M13 sequencing primer sites located in the pGEM-T vector. Sequencing was carried out on an ABI377 automated sequencer (CA USA) as described above, after carrying out the sequencing reaction on an ABI Catalyst (CA USA). Sequence gaps and areas of inadequate coverage were PCR amplified directly from O157 chromosomal DNA using primers based on the already obtained sequencing data and sequenced using the standard M13 sequencing primer sites attached to the PCR primers.

#### D. Analysis of the *E. coli* O157 O antigen gene cluster nucleotide sequence data

Sequence data were processed and analysed using the Staden programs [Staden, R., 1982 "Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing." *Nuc. Acid Res.* 10: 4731-4751; Staden, R., 1986 "The current status and portability of our sequence handling software". *Nuc. Acid Res.* 14: 217-231; Staden, R. 1982 "An interactive graphics program for comparing and aligning nucleic acid and amino acid sequence". *Nuc. Acid Res.* 10: 2951-2961].

Figure 4 shows the structure of *E. coli* O157 O antigen gene cluster. Twelve open reading frames were predicted from the sequence data, and the nucleotide and amino acid sequences of all these genes were then used to search the GenBank database for indication of possible function and

specificity of these genes. The position of each gene is listed in Table 9. The nucleotide sequence is presented in Figure 6.

orfs 10 and 11 showed high level identity to *manC* and *manB* and were named *manC* and *manB* respectively. *orf7* showed 89% identity (at amino acid level) to the *gmd* gene of the *E. coli* colanic acid capsule gene cluster (Stevenson G., K. et al. 1996 "Organisation of the *Escherichia coli* K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid". J. Bacteriol. 178:4885-4893) and was named *gmd*. *orf8* showed 79% and 69% identity (at amino acid level) respectively to *wcaG* of the *E. coli* colanic acid capsule gene cluster and to *wbcJ* (*orf14.8*) gene of the *Yersinia enterocolitica* O8 O antigen gene cluster (Zhang, L. et al. 1997 "Molecular and chemical characterization of the lipopolysaccharide O-antigen and its role in the virulence of *Y. enterocolitica* serotype O8". Mol. Microbiol. 23:63-76). Colanic acid and the *Yersinia* O8 O antigen both contain fucose as does the O157 O antigen. There are two enzymatic steps required for GDP-L-fucose synthesis from GDP-4-keto-6-deoxy-D-mannose, the product of the *gmd* gene product. However, it has been shown recently (Tonetti, M et al. 1996 Synthesis of GDP-L-fucose by the human FX protein J. Biol. Chem. 271:27274-27279) that the human FX protein has "significant homology" with the *wcaG* gene (referred to as *Yefb* in that paper), and that the FX protein carries out both reactions to convert GDP-4-keto-6-deoxy-D-mannose to GDP-L-fucose. We believe that this makes a very strong case for *orf8* carrying out these two steps and propose to name the gene *fcl*. In support of the one enzyme carrying out both functions is the observation that there are no genes other than *manB*, *manC*, *gmd* and *fcl* with similar levels of similarity between the three bacterial gene clusters for fucose containing structures.

*orf5* is very similar to *wbeE* (*rfbE*) of *Vibrio*

*cholerae* O1, which is thought to be the perosamine synthetase, which converts GDP-4-keto-6-deoxy-D-mannose to GDP-perosamine (Stroeher, U.H et al. 1995 "A putative pathway for perosamine biosynthesis is the first function encoded within the *rfb* region of *Vibrio cholerae*" O1. Gene 166: 33-42). *V. cholerae* O1 and *E. coli* O157 O antigens contain perosamine and N-acetyl-perosamine respectively. The *V. cholerae* O1 *manA*, *manB*, *gmd* and *wbeE* genes are the only genes of the *V. cholerae* O1 gene cluster with significant similarity to genes of the *E. coli* O157 gene cluster and we believe that our observations both confirm the prediction made for the function of *wbe* of *V. cholerae*, and show that *orf5* of the O157 gene cluster encodes GDP-perosamine synthetase.

*orf5* is therefore named *per*. *orf5* plus about 100bp of the upstream region (position 4022-5308) was previously sequenced by Bilge, S.S. et al. [1996 "Role of the *Escherichia coli* O157-H7 O side chain in adherence and analysis of an *rfb* locus". *Infect. Immun.* 64:4795-4801].

*orf12* shows high level similarity to the conserved region of about 50 amino acids of various members of an acetyltransferase family (Lin, W., et al. 1994 "Sequence analysis and molecular characterisation of genes required for the biosynthesis of type 1 capsular polysaccharide in *Staphylococcus aureus*". *J. Bacteriol.* 176: 7005-7016) and we believe it is the N-acetyltransferase to convert GDP-perosamine to GDP-perNac. *orf12* has been named *wbdR*.

The genes *manB*, *manC*, *gmd*, *fcl*, *per* and *wbdR* account for all of the expected biosynthetic pathway genes of the O157 gene cluster.

The remaining biosynthetic step(s) required are for synthesis of UDP-GalNac from UDP-Glc. It has been proposed (Zhang, L., et al. 1997 "Molecular and chemical characterisation of the lipopolysaccharide O-antigen and its role in the virulence of *Yersinia enterocolitica* serotype O8". *Mol. Microbiol.* 23:63-76) that in *Yersinia enterocolitica* UDP-GalNac is synthesised from UDP-GlcNac

by a homologue of galactose epimerase (GalE), for which there is a *galE* like gene in the *Yersinia enterocolitica* O8 gene cluster. In the case of O157 there is no *galE* homologue in the gene cluster and it is not clear how UDP-GalNAc is synthesised. It is possible that the galactose epimerase encoded by the *galE* gene in the *gal* operon, can carry out conversion of UDP-GlcNAc to UDP-GalNAc in addition to conversion of UDP-Glc to UDP-Gal. There do not appear to be any gene(s) responsible for UDP-GalNAc synthesis in the O157 gene cluster.

*orf4* shows similarity to many *wzx* genes and is named *wzx* and *orf2* which shows similarity of secondary structure in the predicted protein to other *wzy* genes and is for that reason named *wzy*.

The *orf1*, *orf3* and *orf6* gene products all have characteristics of transferases, and have been named *wbdN*, *wbdO* and *wbdP* respectively. The O157 O antigen has 4 sugars and 4 transferases are expected. The first transferase to act would put a sugar phosphate onto undecaprenol phosphate. The two transferases known to perform this function, WbaP (RfbP) and WecA (Rfe) transfer galactose phosphate and N-acetyl-glucosamine phosphate respectively to undecaprenol phosphate. Neither of these sugars is present in the O157 structure.

Further, none of the presumptive transferases in the O157 gene cluster has the transmembrane segments found in WecA and WbaP which transfer a sugar phosphate to undecaprenol phosphate and expected for any protein which transferred a sugar to undecaprenol phosphate which is embedded within the membrane.

The WecA gene which transfers GlcNAc-P to undecaprenol phosphate is located in the Enterobactereal Common Antigen (ECA) gene cluster and it functions in ECA synthesis in most and perhaps all *E. coli* strains, and also in O antigen synthesis for those strains which have GlcNAc as the first sugar in the O unit.

It appears that WecA acts as the transferase for

addition of GalNAc-1-P to undecaprenol phosphate for the *Yersinia enterocolitica* O8 O antigen [Zhang et al.1997 "Molecular and chemical characterisation of the lipopolysaccharide O antigen and its role in the virulence of *Yersinia enterocolitica* serotype O8" Mol. Microbiol. 23: 63-76.] and perhaps does so here as the O157 structure includes GalNAc. *WecA* has also been reported to add Glucose-1-P phosphate to undecaprenol phosphate in *E. coli* O8 and O9 strains, and an alternative possibility for transfer of the first sugar to undecaprenol phosphate is *WecA* mediated transfer of glucose, as there is a glucose residue in the O157 O antigen. In either case the requisite number of transferase genes are present if GalNAc or Glc is transferred by *WecA* and the side chain Glc is transferred by a transferase outside of the O antigen gene cluster.

*orf9* shows high level similarity (44% identity at amino acid level, same length) with *wcaH* gene of the *E. coli* colanic acid capsule gene cluster. The function of this gene is unknown, and we give *orf9* the name *wbdQ*.

The DNA between *manB* and *wdbR* has strong sequence similarity to one of the H-repeat units of *E. coli* K12. Both of the inverted repeat sequences flanking this region are still recognisable, each with two of the 11 bases being changed. The H-repeat associated protein encoding gene located within this region has a 267 base deletion and mutations in various positions. It seems that the H-repeat unit has been associated with this gene cluster for a long period of time since it translocated to the gene cluster, perhaps playing a role in assembly of the gene cluster as has been proposed in other cases.

#### Materials and Methods - part 4

To test our hypothesis that O antigen genes for transferases and the *wzx*, *wzy* genes were more specific than pathway genes for diagnostic PCR, we first carried out PCR using primers for all the *E. coli* O16 O antigen

genes (Table 7). The PCR was then carried out using PCR primers for *E. coli* 0111 transferase, *wzx* and *wzy* genes (Table 8, 8A). PCR was also carried out using PCR primers for the *E. coli* 0157 transferase, *wzx* and *wzy* genes (Table 9, 9A).

Chromosomal DNA from the 166 serotypes of *E. coli* available from Statens Serum Institut, 5 Artillerivej, 2300 Copenhagen Denmark was isolated using the Promega Genomic (Madison WI USA) isolation kit. Note that 164 of the serogroups are described by Ewing W. H.: Edwards and Elwings "Identification of the Enterobacteriaceae" Elsevier, Amsterdam 1986 and that they are numbered 1-171 with numbers 31, 47, 67, 72, 93, 94 and 122 no longer valid. Of the two serogroup 19 strains we used 19ab strain F8188-41. Lior H. 1994 ["Classification of *Escherichia coli* In *Escherichia coli* in domestic animals and humans pp 31-72. Edited by C.L. Gyles CAB international] adds two more numbered 172 and 173 to give the 166 serogroups used. Pools containing 5 to 8 samples of DNA per pool were made. Pool numbers 1 to 19 (Table 4) were used in the *E. coli* 0111 and 0157 assay. Pool numbers 20 to 28 were also used in the 0111 assay, and pool numbers 22 to 24 contained *E. coli* 0111 DNA and were used as positive controls (Table 5). Pool numbers 29 to 42 were also used in the 0157 assay, and pool numbers 31 to 36 contained *E. coli* 0157 DNA, and were used as positive controls (Table 6). Pool numbers 2 to 20, 30, 43 and 44 were used in the *E. coli* 016 assay (Tables 4 to 6). Pool number 44 contained DNA of *E. coli* K-12 strains C600 and WG1 and was used as a positive control as between them they have all of the *E. coli* K-12 016 O antigen genes.

PCR reactions were carried out under the following conditions: denaturing 94°C/30"; annealing, temperature varies (refer to Tables)/30"; extension, 72°C/1'; 30 cycles. PCR reaction was carried out in an volume of 25mL for each pool. After the PCR reaction, 10mL PCR



product from each pool was run on an agarose gel to check for amplified DNA.

Each *E. coli* chromosomal DNA sample was checked by gel electrophoresis for the presence of chromosomal DNA and by PCR amplification of the *E. coli mdh* gene using oligonucleotides based on *E. coli* K-12 [Boyd et al. (1994) "Molecular genetic basis of allelic polymorphism in malate dehydrogenase (*mdh*) in natural populations of *Escherichia coli* and *Salmonella enterica*" Proc. Nat. Acad. Sci. USA. 91:1280-1284.] Chromosomal DNA samples from other bacteria were only checked by gel electrophoresis of chromosomal DNA.

A. Primers based on *E. coli* O16 O antigen gene cluster sequence.

The O antigen gene cluster of *E. coli* O16 was the only typical *E. coli* O antigen gene cluster that had been fully sequenced prior to that of O111, and we chose it for testing our hypothesis. One pair of primers for each gene was tested against pools 2 to 20, 30 and 43 of *E. coli* chromosomal DNA. The primers, annealing temperatures and functional information for each gene are listed in Table 8.

For the five pathway genes, there were 17/21, 13/21, 0/21, 0/21, 0/21 positive pools for *xmlB*, *xmlD*, *xmlA*, *xmlC* and *glf* respectively (Table 7). For the *wzx*, *wzy* and three transferase genes there were no positives amongst the 21 pools of *E. coli* chromosomal DNA tested (Table 7). In each case the #44 pool gave a positive result.

B. Primers based on the *E. coli* O111 O antigen gene cluster sequence.

One to four pairs of primers for each of the transferase, *wzx* and *wzy* genes of O111 were tested against the pools 1 to 21 of *E. coli* chromosomal DNA (Table 8). For *wbdH*, four pairs of primers, which bind

to various regions of this gene, were tested and found to be specific for O111 as there was no amplified DNA of the correct size in any of those 21 pools of *E. coli* chromosomal DNA tested. Three pairs of primers for *wbdM* were tested, and they are all specific although primers #985/#986 produced a band of the wrong size from one pool. Three pairs of primers for *wzx* were tested and they all were specific. Two pairs of primers were tested for *wzy*, both are specific although #980/#983 gave a band of the wrong size in all pools. One pair of primers for *wbdL* was tested and found unspecific and therefore no further test was carried out. Thus, *wzx*, *wzy* and two of the three transferase genes are highly specific to O111.

Bands of the wrong size found in amplified DNA are assumed to be due to chance hybridisation of genes widely present in *E. coli*. The primers, annealing temperatures and positions for each gene are in Table 8.

The O111 assay was also performed using pools including DNA from O antigen expressing *Yersinia pseudotuberculosis*, *Shigella boydii* and *Salmonella enterica* strains (Table 8A). None of the oligonucleotides derived from *wbdH*, *wzx*, *wzy* or *wbdM* gave amplified DNA of the correct size with these pools. Notably, pool number 25 includes *S. enterica* Adelaide which has the same O antigen as *E. coli* O111: this pool did not give a positive PCR result for any primers tested indicating that these genes are highly specific for *E. coli* O111.

Each of the 12 pairs binding to *wbdH*, *wzx*, *wzy* and *wbdM* produces a band of predicted size with the pools containing O111 DNA (pools number 22 to 24). As pools 22 to 24 included DNA from all strains present in pool 21 plus O111 strain DNA (Table 5), we conclude that the 12 pairs of primers all give a positive PCR test with each of three unrelated O111 strains but not with any other strains tested. Thus these genes are highly specific for *E. coli* O111.

C. Primers based on the *E. coli* 0157 O antigen gene cluster sequence.

Two or three primer pairs for each of the transferase, *wzx* and *wzy* genes of 0157 were tested against *E. coli* chromosomal DNA of pools 1 to 19, 29 and 30 (Table 9). For *wbdN*, three pairs of primers, which bind to various regions of this gene, were tested and found to be specific for 0157 as there was no amplified DNA in any of those 21 pools of *E. coli* chromosomal DNA tested. Three pairs of primers for *wbdO* were tested, and they are all specific although primers # 1211/#1212 produced two or three bands of the wrong size from all pools. Three pairs of primers were tested for *wbdP* and they all were specific. Two pairs of primers were tested for *wbdR* and they were all specific. For *wzy*, three pairs of primers were tested and all were specific although primer pair #1203/#1204 produced one or three bands of the wrong size in each pool. For *wzx*, two pairs of primers were tested and both were specific although primer pair #1217/#1218 produced 2 bands of wrong size in 2 pools, and 1 band of wrong size in 7 pools. Bands of the wrong size found in amplified DNA are assumed to be due to chance hybridisation of genes widely present in *E. coli*. The primers, annealing temperatures and function information for each gene are in Table 9.

The 0157 assay was also performed using pools 37 to 42, including DNA from O antigen expressing *Yersinia pseudotuberculosis*, *Shigella boydii*, *Yersinia enterocolitica* 09, *Brucella abortus* and *Salmonella enterica* strains (Table 9A). None of the oligonucleotides derived from *wbdN*, *wzy*, *wbdO*, *wzx*, *wbdP* or *wbdR* reacted specifically with these pools, except that primer pair #1203/#1204 produced two bands with *Y. enterocolitica* 09 and one of the bands is of the same size with that from the positive control. Primer pair #1203/#1204 binds to *wzy*. The predicted secondary

structures of Wzy proteins are generally similar, although there is very low similarity at amino acid or DNA level among the sequenced wzy genes. Thus, it is possible that *Y. enterocolitica* O9 has a wzy gene closely related to that of *E. coli* O157. It is also possible that this band is due to chance hybridization of another gene, as the other two wzy primer pairs (#1205/#1206 and #1207/#1208) did not produce any band with *Y. enterocolitica* O9. Notably, pool number 37 includes *S. enterica* Landau which has the same O antigen as *E. coli* O157, and pool 38 and 39 contain DNA of *B. abortus* and *Y. enterocolitica* O9 which cross react serologically with *E. coli* O157. This result indicates that these genes are highly O157 specific, although one primer pair may have cross reacted with *Y. enterocolitica* O9.

Each of the 16 pairs binding to *wbdN*, *wzx*, *wzy*, *wbdO*, *wbdP* and *wbdR* produces a band of predicted size with the pools containing O157 DNA (pools number 31 to 36). As pool 29 included DNA from all strains present in pools 31 to 36 other than O157 strain DNA (Table 6), we conclude that the 16 pairs of primers all give a positive PCR test with each of the five unrelated O157 strains.

Thus PCR using primers based on genes *wbdN*, *wzy*, *wbdO*, *wzx*, *wbdP* and *wbdR* is highly specific for *E. coli* O157, giving positive results with each of six unrelated O157 strains while only one primer pair gave a band of the expected size with one of three strains with O antigens known to cross-react serologically with *E. coli* O157.

TABLE 1

H7 strains used in this work in addition to the H  
antigens type strains

5

Name used in this study	Serotype	Original name	Source*
M527	O157:H7	C664-1992	a
M917	O18ac:H7	A57	IMVS
M918	O18ac:H7	A62	IMVS
M973	O2:H7	A1107	CDC
M1004	O157:H7	EH7	b
M1179	O18ac:H7	D-M3291/54	IMVS
M1200	O7:H7	A64	c
M1211	O19ab:H7	F8188-41	IMVS
M1328	O53:H7	14097	IMVS
M1686	O55:H7	TB156	d

\*

10

a. Statens Serum Institut, Copenhagen, Denmark.

b. Dr R. Brown of Royal Children's Hospital, Melbourne, Australia.

15

c. Max-Planck Institut fur molekulare Genetik, Berlin, Germany.

d. Dr P. Tarr of Children's Hospital and Medical Center, University of Washington, USA.

20

IMVS, Institute of Medical and veterinary Science, Adelaide, Australia.

CDC, Centers for Disease Control and prevention, Atlanta, USA.

**Table 2**  
**Oligonucleotides used to PCR amplify *fliC* genes**  
**from different H type strains for sequencing**

H Type Strains	Annealing Temperature (°C)	Primers Used
1	55	#1575/#1576
2	55	#1285/#1286
3	55	#1285/#1286
4	50	S1431/#1432
5	60	#1285/#1286
6	55	#1575/#1576
7	55	#1575/#1576
8	55	#1431/#1432
9	60	#1575/#1576
10	55	#1575/#1576
11	55	#1285/#1286
12	60	#1575/#1576
14	60	#1575/#1576
15	60	#1575/#1576
16	60	#1575/#1576
17	60	#1417/#1418
18	60	#1575/#1576
19	60	#1575/#1576
20	60	#1575/#1576
21	55	#1285/#1286
23	60	#1575/#1576
24	60	#1285/#1286
25	60	#1417/#1418
26	60	#1575/#1576
27	50	#1431/#1432
28	60	#1575/#1576
29	60	#1285/#1286
30	60	#1575/#1576
31	60	#1575/#1576
32	60	#1575/#1576
33	60	#1285/#1286
34	55	#1575/#1576
35	50	#1431/#1432
37	60	#1285/#1286
38	60	#1285/#1286
39	55	#1285/#1286
40	55	#1285/#1286
41	60	#1575/#1576
42	60	#1285/#1286
43	60	#1575/#1576
44	60	#1285/#1286
45	60	#1575/#1576
46	60	#1575/#1576
47	55	#1285/#1286
48	60	#1575/#1576
49	60	#1575/#1576
50	60	#1285/#1286
51	60	#1575/#1576
52	60	#1575/#1576
54	50	#1431/#1432
55	60	#1285/#1286
56	60	#1285/#1286

**Table 3**  
**Summary of the flagellin sequences obtained and specific H type**  
**oligonucleotide primers**

H type strain(s) the sequenced gene(s) obtained from	H specificity coded by the gene(s)	H type strain from which the flagellin gene sequence was used for primer choice	Positions of primer 1	Positions of primer 2
1	1	1	892-909	1172-1189
2	2	2	568-587	1039-1056
4,17,44	4	4	466-483	628-648
5	5	5	697-714	877-897
6	6	6	565-585	799-816
7	7	7	553-570 (primer #1806)	1483-1500 (primer #1809)
9	9	9	616-633	838-855
10(50)***	10	10	559-579	697-717
11	11	11	586-606*	791-810*
12	12	12	892-909	1172-1189
14	14	14	586-606	793-813
15	15	15	640-660	817-834
3	16	3	649-666	925-942
18	18	18	589-606	802-819
19	19	19	607-624	538-555
20	20	20	574-591	760-780
21,47	21	21	676-693**	862-879**
23	23	23	637-654	1336-1353
24	24	24	496-516	772-792
26	26	26	553-570	772-789
27	27	27	685-702	799-819
28	28	28	592-609	778-798
29	29	29	538-555	757-774
30	30	30	814-831	943-962
31	31	31	571-588	790-807
32	32	32	514-831	1057-1074
33	33	33	553-570	718-735
34	34	34	568-585	796-816
38,55	38	38	553-573	709-729
39	39	39	556-573	718-735
41	41	41	598-615	784-801
42	42	42	547-567	715-735
43	43	43	580-597	844-861
45	45	45	640-657	943-963
46	46	46	565-582	781-801
49	49	49	589-609	754-771
51	51	51	565-582	1042-1059
52	52	52	598-615	829-846
56	56	56	697-714	877-897
8 and 40		8	562-579	1045-1062
25		25	529-549	703-723
35		non-functional H11 gene	769-789*	1045-1065*
37		37	520-537	715-735
48		48	568-585	835-852
54		non-functional H21 gene	988-1008**	1344-1364**

\* See section 13 for choice of primers for the flagellin gene of H11  
 \*\* See section 13 for choice of primers for the flagellin gene of H21  
 \*\*\* See text

**Table 3A**  
**Cloning, expression and identification of flagellin genes**

H type strain from which the H antigen gene was amplified	Primers used for PCR amplification of the H antigen gene	Annealing temperature (°C) used for PCR amplification	Plasmid carrying the H antigen gene	Host strain used for expression	Anti-serum which reacts with an <i>E. Coli</i> <i>flhC</i> deletion strain carrying the plasmid	H antigen encoded by the cloned gene
H1	#1868 & #1870	55	pPR1920	M2126	H1	H1
H2	#1868 & #1870	55	pPR1977	P5560	H2	H2
H3	#1868 & #1870	55	pPR1969	P5560	H16	H16
H4	#1878 & #1885	65	pPR1955	P5560	H4	H4
H5	#1868 & #1870	60	pPR1967	M2126	H5	H5
H6	#1868 & #1870	55	pPR1921	P5560	H6	H6
H7	#1868 & #1870	55	pPR1919	P5560	H7	H7
H9	#1868 & #1870	55	pPR1922	P5560	H9	H9
H10	#1868 & #1870	55	pPR1923	P5560	H10	H10
H11	#1868 & #1870	55	pPR1981	M2126	H11	H11
H12	#1868 & #1870	60	pPR1990	M2126	H12	H12
H14	#1868 & #1870	55	pPR1924	P5560	H14	H14
H15	#1868 & #1870	55	pPR1925	P5560	H15	H15
H17	#1878 & #1885	65	pPR1957	P5560	H4	H4
H18	#1868 & #1870	55	pPR1986	M2126	H18	H18
H19	#1868 & #1870	55	pPR1927	P5560	H19	H19
H20	#1868 & #1870	55	pPR1963	M2126	H20	H20
H21	#1868 & #1870	55	pPR1995	M2126	H21	H21
H23	#1868 & #1869	55	pPR1942	P5560	H23	H23
H24	#1868 & #1870	55	pPR1971	M2126	H24	H24
H26	#1868 & #1870	65	pPR1928	P5560	H26	H26
H27	#1868 & #1870	55	pPR1970	M2126	H27	H27
H28	#1868 & #1870	60	pPR1944	P5560	H28	H28
H29	#1868 & #1870	55	pPR1972	M2126	H29	H29
H30	#1868 & #1871	55	pPR1948	P5560	H30	H30
H31	#1868 & #1870	65	pPR1965	M2126	H31	H31
H32	#1868 & #1871	55	pPR1940	P5560	H32	H32
H33	#1868 & #1871	55	pPR1976	M2126	H33	H33
H34	#1868 & #1870	65	pPR1930	P5560	H34	H34
H38	#1868 & #1870	48	pPR1984	M2126	H38	H38
H39	#1868 & #1870	48	pPR1982	M2126	H39	H39
H41	#1868 & #1870	65	pPR1931	P5560	H41	H41
H42	#1868 & #1870	50	pPR1979	M2126	H42	H42
H43	#1868 & #1870	65	pPR1968	M2126	H43	H43
H45	#1868 & #1870	60	pPR1943	P5560	H45	H45
H46	#1868 & #1870	60	pPR1966	M2126	H46	H46
H49	#1868 & #1870	60	pPR1985	M2126	H49	H49
H51	#1868 & #1870	65	pPR1941	P5560	H51	H51
H52	#1868 & #1870	65	pPR1935	P5560	H52	H52
H56	#1868 & #1870	50	pPR1978	M2126	H56	H56



**Table 3B** Oligonucleotide primers used for PCR amplification and cloning of H antigen genes

#1868 5'- cat gcc atg gca caa gtc att aat acc -3'  
*NcoI*

#1869 5'- ata tgt cga ctt aac cct gca gca gag aca g -3'  
*Sall*

#1870 5' - atg gat cct taa ccc tgc agc aga gac ag -3'  
*BamHI*

#1871 5' - aac tgc agt taa ccc tgt agc aga gac ag -3'  
*PstI*

#1872 5' - cgg gat ccc gca gac tgg ttc ttg ttg at - 3'  
*BamHI*

#1878 5' - cgg gat cca ctt cta tcg agc gcc tct ct - 3'  
*BamHI*

#1884 5' - gct cta gag cgc aga tca ttc agc agg cc -3'  
*XbaI*

#1885 5' - gct cta gac atg ttg gac act tcg gtc gc - 3'  
*XbaI*

- 75 -

TABLE 4

Pool No.	Strains of which chromosomal DNA included in the pool	Source*
1	<i>E. coli</i> type strains for O serotypes 1, 2, 3, 4, 10, 16, 18 and 39	IMVS <sup>a</sup>
2	<i>E. coli</i> type strains for O serotypes 40, 41, 48, 49, 71, 73, 88 and 100	IMVS
3	<i>E. coli</i> type strains for O serotypes 102, 109, 119, 120, 121, 125, 126 and 137	IMVS
4	<i>E. coli</i> type strains for O serotypes 138, 139, 149, 7, 5, 6, 11 and 12	IMVS
5	<i>E. coli</i> type strains for O serotypes 13, 14, 15, 17, 19ab, 20, 21 and 22	IMVS
6	<i>E. coli</i> type strains for O serotypes 23, 24, 25, 26, 27, 28, 29 and 30	IMVS
7	<i>E. coli</i> type strains for O serotypes 32, 33, 34, 35, 36, 37, 38 and 42	IMVS
8	<i>E. coli</i> type strains for O serotypes 43, 44, 45, 46, 50, 51, 52 and 53	IMVS
9	<i>E. coli</i> type strains for O serotypes 54, 55, 56, 57, 58, 59, 60 and 61	IMVS
10	<i>E. coli</i> type strains for O serotypes 62, 63, 64, 65, 66, 68, 69 and 70	IMVS
11	<i>E. coli</i> type strains for O serotypes 74, 75, 76, 77, 78, 79, 80 and 81	IMVS
12	<i>E. coli</i> type strains for O serotypes 82, 83, 84, 85, 86, 87, 89 and 90	IMVS
13	<i>E. coli</i> type strains for O serotypes 91, 92, 95, 96, 97, 98, 99 and 101	IMVS
14	<i>E. coli</i> type strains for O serotypes 103, 104, 105, 106, 107, 108 and 110	IMVS
15	<i>E. coli</i> type strains for O serotypes 112, 162, 113, 114, 115, 116, 117 and 118	IMVS
16	<i>E. coli</i> type strains for O serotypes 123, 165, 166, 167, 168, 169, 170 and 171	See b
17	<i>E. coli</i> type strains for O serotypes 172, 173, 127, 128, 129, 130, 131 and 132	See c
18	<i>E. coli</i> type strains for O serotypes 133, 134, 135, 136, 140, 141, 142 and 143	IMVS
19	<i>E. coli</i> type strains for O serotypes 144, 145, 146, 147, 148, 150, 151 and 152	IMVS

\*

- a. Institute of Medical and Veterinary Science, Adelaide, Australia
- b. 123 from IMVS; the rest from Statens Serum Institut, Copenhagen, Denmark
- c. 172 and 173 from Statens Serum Institut, Copenhagen, Denmark, the rest from IMVS

TABLE 5

Pool No.	Strains of which chromosomal DNA included in the pool	Source*
20	<i>E. coli</i> type strains for O serotypes 153, 154, 155, 156, 157, 158, 159 and 160	IMVS
21	<i>E. coli</i> type strains for O serotypes 161, 163, 164, 8, 9 and 124	IMVS
22	As pool #21, plus <i>E. coli</i> 0111 type strain Stoke W.	IMVS
23	As pool #21, plus <i>E. coli</i> 0111:H2 strain C1250-1991	See d
24	As pool #21, plus <i>E. coli</i> 0111:H12 strain C156-1989	See e
25	As pool #21, plus <i>S. enterica</i> serovar Adelaide	See f
26	<i>Y. pseudotuberculosis</i> strains of O groups IA, IIA, IIB, IIC, III, IVA, IVB, VA, VB, VI and VII	See g
27	<i>S. boydii</i> strains of serogroups 1, 3, 4, 5, 6, 8, 9, 10, 11, 12, 14 and 15	See h
28	<i>S. enterica</i> strains of serovars (each representing a different O group) Typhi, Montevideo, Ferruch, Jangwani, Raus, Hvittingfoss, Waycross, Dan, Dugbe, Basel, 65:i:e,n,z,15 and 52:d:e,n,x,z,15	IMVS

\*

- d. C1250-1991 from Statens Serum Institut, Copenhagen, Denmark
- e. C156-1989 from Statens Serum Institut, Copenhagen, Denmark
- f. *S. enterica* serovar Adelaide from IMVS
- g. Dr S Aleksic of Institute of Hygiene, Germany
- h. Dr J Lefebvre of Bacterial Identification Section, Laboratoire de Santé Publique du Québec, Canada

TABLE 6

Pool No.	Strains of which chromosomal DNA included in the pool	Source*
29	<i>E. coli</i> type strains for O serotypes 153, 154, 155, 156, 158, 159 and 160	IMVS
30	<i>E. coli</i> type strains for O serotypes 161, 163, 164, 8, 9, 111 and 124	IMVS
31	As pool #29, plus <i>E. coli</i> O157 type strain A2 (O157:H19)	IMVS
32	As pool #29, plus <i>E. coli</i> O157:H16 strain C475-89	See d
33	As pool #29, plus <i>E. coli</i> O157:H45 strain C727-89	See d
34	As pool #29, plus <i>E. coli</i> O157:H2 strain C252-94	See d
35	As pool #29, plus <i>E. coli</i> O157:H39 strain C258-94	See d
36	As pool #29, plus <i>E. coli</i> O157:H26	See e
37	As pool #29, plus <i>S. enterica</i> serovar Landau	See f
38	As pool #29, plus <i>Brucella abortus</i>	See g See h
39	As pool #29, plus <i>Y. enterocolitica</i> O9	
40	<i>Y. pseudotuberculosis</i> strains of O groups IA, IIA, IIB, IIC, IIL, IVA, IVB, VA, VB, VI and VII	See i
41	<i>S. boydii</i> strains of serogroups 1, 3, 4, 5, 6, 8, 9, 10, 11, 12, 14 and 15	See j
42	<i>S. enterica</i> strains of serovars (each representing a different O group) Typhi, Montevideo, Ferruch, Jangwani, Raus, Hvittingfoss, Waycross, Dan, Dugbe, Basel, 65:i:e,n,z15 and 52:d:e,n,x,z15	IMVS
43	<i>E. coli</i> type strains for O serotypes 1,2,3,4,10,18 and 29	IMVS
44	As pool #43, plus <i>E. coli</i> K-12 strains C600 and WG1	IVMS See k

\*

- d. O157 strains from Statens Serum Institut, Copenhagen, Denmark
- e. O157:H26 from Dr R Brown of Royal Children's Hospital, Melbourne, Victoria
- f. *S. enterica* serovar Landau from Dr M Poppoff of Institut Pasteur, Paris, France
- g. *B. Abortus* from the culture collection of The University of Sydney, Sydney, Australia
- h. *Y. enterocolitica* O9 from Dr. K. Bettelheim of Victorian Infectious Diseases Reference Laboratory Victoria, Australia.
- i. Dr S Aleksic of Institute of Hygiene, Germany
- j. Dr J Lefebvre of Bacterial Identification Section, Laboratoire de Santé Publique du Québec, Canada
- k. Strains C600 and WG1 from Dr. B.J. Backmann of Department of Biology, Yale University, USA.

TABLE 7 PCR assay result using primers based on the *E. coli* serotype O16 (strain K-12) O antigen gene cluster sequence

Gene	Function	Base positions of the gene	Forward primer (base positions)	Reverse primer (base positions)	Length of the PCR fragment	Number of pools (out of 21) giving band of correct size	Annealing temperature of the PCR
<i>rmlB</i> *	TDP-ribose pathway	90-1175	#1064(91-109)	#1065(1175-1157)	1085bp	17	60°C
<i>rmlD</i> *	TDP-ribose pathway	1175-2074	#1066(1175-1193)	#1067(2075-2058)	901bp	13	60°C
<i>rmlA</i> *	TDP-ribose pathway	2132-3013	#1068(2131-2148)	#1069(3013-2995)	883bp	0	60°C
<i>rmlC</i> *	TDP-ribose pathway	3013-3570	#1070(3012-3029)	#1071(3570-3551)	559bp	0	60°C
<i>glt</i> *	Galactofuranose pathway	4822-5925	#1074(4822-4840)	#1075(5925-5908)	1104bp	0	55°C
<i>uzx</i> *	Flippase	3567-4814	#1072(3567-3586)	#1073(4814-4797)	1248bp	0	55°C
<i>ury</i> *	O polymerase	5925-7091	#1076(5925-5944)	#1077(7091-7074)	1167bp	0	60°C
<i>ubt1</i> *	Galactofuranosyl transferase	7094-8086	#1078(7094-7111)	#1079(8086-8069)	993bp	0	50°C
<i>ubt2</i> *	Acetyltransferase	8067-8654	#1080(8067-8084)	#1081(8654-8632)	588bp	0	60°C
<i>ubt3</i> **	Glucosyl transferase	5770-6888	#1082(5770-5787)	#1083(6888-6871)	1119bp	0	55°C
<i>ubt4</i> ***	Rhamnosyltransferase	679-1437	#1084(679-697)	#1085(1473-1456)	795bp	0***	55°C

\*, \*\*, \*\*\* Base positions based on GenBank entry U09876, U03041 and L19537 respectively  
 19 pools giving a band of wrong size

\*\*\*\*

TABLE 8 PCR assay data using 0111 primers

Gene	Base positions of the gene according to SEQ ID NO: 1	Forward primer (base positions)	Reverse primer (base positions)	Length of the PCR fragment	Number of pools (out of 21) giving band of correct size	Annealing temperature of the PCR
<i>wbdH</i>	739-1932	#866 (739-757)	#867 (1941-1924)	1203bp	0	60°C
		#976 (925-942)	#978 (1731-1714)	807bp	0	60°C
		#976 (925-942)	#979 (1347-1330)	423bp	0	60°C
		#977 (1165-1182)	#978 (1731-1714)	567bp	0	60°C
<i>wzx</i>	8646-9911	#969 (8646-8663)	#970 (9908-9891)	1263bp	0	50°C
		#1060 (8906-8923)	#1062 (9468-9451)	563bp	0	60°C
		#1061 (9150-9167)	#1063 (9754-9737)	605bp	0	50°C
<i>wzy</i>	9901-10953	#900 (9976-9996)	#901 (10827-10807)	852bp	0	60°C
		#980 (10113-10130)	#983 (10484-10467)	372bp	0*	61°C
<i>wbdL</i>	10931-11824	#870 (10931-10949)	#871 (11824-11796)	894bp	7	60°C
<i>wbdM</i>	11821-12945	#868 (11821-11844)	#869 (12945-12924)	1125bp	0	60°C
		#984 (12042-12059)	#987 (12447-12430)	406bp	0	60°C
		#985 (12258-12275)	#986 (12698-12681)	441bp	0**	65°C

\* Giving a band of wrong size in all pools

\*\* One pool giving a band of wrong size

TABLE 8A PCR specificity test data using 0111 primers

Gene	Base positions of the gene according to SEQ ID NO: 1	Forward primer (base positions)	Reverse primer (base positions)	Length of the PCR fragment	Number of pools (pools no. 25-28) giving band of correct size	Annealing temperature of the PCR
<i>tubH</i>	739-1932	#866 (739-757)	#867 (1941-1924)	1203bp	0*	60°C
		#976 (925-942)	#978 (1731-1714)	807bp	0	60°C
		#976 (925-942)	#979 (1347-1330)	423bp	0	60°C
		#977 (1165-1182)	#978 (1731-1714)	567bp	0	60°C
<i>uizx</i>	8646-9911	#969 (8646-8663)	#970 (9908-9891)	1263bp	0	55°C
		#1060 (8906-8923)	#1062 (9468-9451)	563bp	0	60°C
		#1061 (9150-9167)	#1063 (9754-9737)	605bp	0*	50°C
<i>uizy</i>	9901-10953	#900 (9976-9996)	#901 (10827-10807)	852bp	0	60°C
		#980 (10113-10130)	#983 (10484-10467)	372bp	0**	60°C
<i>tubL</i>	10931-11824	#870 (10931-10949)	#871 (11824-11796)	894bp	0	60°C
<i>tubM</i>	11821-12945	#868 (11821-11844)	#869 (12945-12924)	1125bp	0	60°C
		#984 (12042-12059)	#987 (12447-12430)	406bp	0	60°C
		#985 (12258-12275)	#986 (12698-12681)	441bp	0*	65°C

\* 1 pool giving a band of wrong size

\*\* 2 pools giving 3 bands of wrong sizes, 1 pool giving 2 bands of wrong sizes

TABLE 9 PCR results using primers based on the *E. coli* O157 sequence

Gene	Function	Base position of the gene according to SEQ ID NO: 2	Forward primer (base positions)	Reverse primer (base positions)	Length of the PCR fragment	Number of pools (out of 21) giving band of correct size	Annealing temperature of the PCR
<i>wbdN</i>	Sugar transferase	79-361	#1197(79-96)	#1198 (861-844)	783	0	55°C
			#1199(184-201)	#1200(531-514)	348	0	55°C
			#1201(310-327)	#1202(768-751)	459	0	55°C
<i>uzy</i>	O antigen	858-2042	#1203(858-875)	#1204(2042-2025)	1185	0*	50°C
			#1205(1053-1070)	#1206(1619-1602)	567	0	63°C
			#1207(1278-1295)	#1208(1913-1896)	636	0	60°C
<i>wbdO</i>	Sugar transferase	2011-2757	#1209(2011-2028)	#1210(2757-2740)	747	0	50°C
			#1211(2110-2127)	#1212(2493-2476)	384	0**	62°C
			#1213(2305-2322)	#1214(2682-2665)	378	0	60°C
<i>uzy</i>	O antigen flippase	2744-4135	#1215(2744-2761)	#1216(4135-4118)	1392	0	50°C
			#1217(2942-2959)	#1218(3638-3611)	687	0***	63°C
<i>wbdP</i>	Sugar transferase	5257-6471	#1221(5257-5274)	#1222(6471-6454)	1215	0	55°C
			#1223(5440-5457)	#1224(5973-5956)	534	0	55°C
			#1225(5707-5724)	#1226(6231-6214)	525	0	55°C
<i>wbdR</i>	N-acetyl	13156-13821	#1229(13261-13278)	#1230(13629-13612)	369	0	55°C
			#1231(13384-13401)	#1232(13731-13714)	348	0	60°C

\* 3 bands of wrong size in one pool, 1 band of wrong size in all other pools

\*\* 3 bands of wrong sizes in 9 pools, 2 bands of wrong size in all other pools

\*\*\* 2 bands of wrong sizes in 2 pools, 1 band of wrong size in 7 pools



TABLE 9A PCR results using primers based on the *E. coli* O157 sequence

Gene	Function	Base position of the gene according to SEQ ID NO: 2	Forward primer (base positions)	Reverse primer (base positions)	Length of the PCR fragment	Number of pools (pools no. 37-42) giving band of correct size	Annealing temperature of the PCR
<i>tubdN</i>	Sugar transferase	79-861	#1197(79-96) #1199(184-201) #1201(310-327)	#1198 (861-844) #1200(531-514) #1202(768-751)	783 348 459	0* 0* 0	55°C 55°C 61°C
<i>uzy</i>	O antigen	858-2042	#1203(858-875) #1205(1053-1070) #1207(1278-1295) #1209(2011-2028)	#1204(2042-2025) #1206(1619-1602) #1208(1913-1896) #1210(2757-2740)	1185 567 636 747	1** 0*** 0 0	50°C 60°C 60°C 50°C
<i>tubdO</i>	Sugar transferase	2011-2757	#1211(2110-2127) #1213(2305-2322) #1215(2744-2761) #1217(2942-2959)	#1212(2493-2476) #1214(2682-2665) #1216(4135-4118) #1218(3628-3611)	384 378 1392 687	0**** 0 0 0	61°C 60°C 50°C 63°C
<i>tubdP</i>	Sugar transferase	5257-6471	#1221(5257-5274) #1223(5440-5457) #1225(5707-5724) #1229(13261-13278) #1231(13384-13401)	#1222(6471-6454) #1224(5973-5956) #1226(6231-6214) #1230(13629-13612) #1233(13731-13714)	1215 534 525 369 348	0 0* 0 0 0	55°C 60°C 55°C 50°C 60°C
<i>tubdR</i>	N-acetyl transferase	13156-13821					

\* 1 band of wrong size in one pool

\*\* pool #39 giving two bands, one band of correct size, the other band of wrong size in another pool.

\*\*\* 2 bands of wrong sizes in one pool

\*\*\*\* 3 bands of wrong sizes in 2 pools, 2 bands of wrong sizes in 2 other pools

- 83 -

CLAIMS:

1. A nucleic acid molecule which encodes all or part of an *E. coli* flagellin protein, the molecule being capable of identifying the H serotype of an *E. coli* when hybridised to a gene of the *E. coli* which encodes a flagellin protein, provided that the molecule does not encode a flagellin protein expressed by the *E. coli* H1, H7, H12 or H48 type strains.
2. A nucleic acid molecule according to claim 1 wherein the molecule is derived from a *fliC* gene.
3. A nucleic acid molecule according to claim 1 including all or part of a sequence according to any one of SEQ ID NOS:1 to 68.
4. A nucleic acid molecule according to claim 1 consisting of all or part of a sequence according to any one of SEQ ID NOS: 1 to 68.
5. A nucleic acid molecule according to claim 4 wherein the molecule is from about 10 to 20 nucleotides in length.
6. A primer selected from the group of primers shown in Table 3.
7. A method of detecting the H serotype of *E. coli* in a sample, the method comprising the following steps:
  - (a) contacting a gene of an *E. coli* in the sample with a nucleic acid molecule according to claim 1 in conditions sufficient to allow the nucleic acid molecule to hybridise to the gene; and
  - (b) detecting a nucleic acid molecule which is hybridised to the gene, to detect the H serotype of the *E. coli* in the sample.

8. A method according to claim 7 wherein the hybridised nucleic acid molecules are detected by Southern Blot analysis.

5

9. A method of detecting the H serotype of *E. coli* in a sample, the method comprising the following steps:

(a) contacting a gene of an *E. coli* in the sample with a pair of nucleic acid molecules according to claim 1 in conditions sufficient to allow the pair of nucleic acid molecules to hybridise to the gene; and

10

(b) detecting a pair of nucleic acid molecules which is hybridised to the gene, to detect the H serotype of the *E. coli* in the sample.

15

10. A method according to claim 9 wherein the hybridised pairs of nucleic acid molecules are detected by the polymerase chain reaction.

20

11. A method for detecting the H and O serotype of *E. coli* in a sample, the method comprising the following steps:

25

(a) contacting a gene of the *E. coli* with a nucleic acid molecule derived from a gene encoding a transferase or a gene encoding an enzyme for the transport or processing of a polysaccharide or oligosaccharide unit, the gene being involved in the synthesis of a *E. coli* O antigen, in conditions sufficient to allow the nucleic acid molecule to hybridise to the gene;

30

(b) contacting a gene of an *E. coli* in the sample with a nucleic acid molecule according to claim 1 in conditions sufficient to allow the nucleic acid molecule to hybridise to the gene; and

35

(c) detecting nucleic acid molecules which are hybridised to the genes, to detect the H and O serotype of the *E. coli* in the sample.

- 85 -

12. A method according to claim 11 wherein the nucleic acid molecule of step (a) is selected from the group consisting of:

- 5      *wbdH* (nucleotide position 739 to 1932 of Figure 5),  
      *wzx* (nucleotide position 8646 to 9911 of Figure 5),  
      *wzy* (nucleotide position 9901 to 10953 of Figure 5),  
      *wbdM* (nucleotide position 11821 to 12945 of Figure 5),  
      *wbdN* (nucleotide position 79 to 861 of Figure 6),  
      *wbdO* (nucleotide position 2011 to 2757 of Figure 6),  
10     *wbdP* (nucleotide position 5257 to 6471 of Figure 6),  
      *wbdR* (nucleotide position 13156 to 13821 of Figure 6),  
      *wzx* (nucleotide position 2744 to 4135 of Figure 6) and  
      *wzy* (nucleotide position 858 to 2042 of Figure 6).

- 15     13. A method according to claim 12 wherein the nucleic acid molecule of step (a) is a primer selected from the group of primers shown in Tables 8, 8A, 9 and 9A.

- 20     14. A method according to claim 11 wherein the hybridised nucleic acid molecules are detected by Southern Blot analysis.

- 25     15. A method for detecting the H and O serotype of *E. coli* in a sample, the method comprising the following steps:

- (a) contacting a gene of the *E. coli* with a pair of nucleic acid molecules derived from a gene encoding a transferase or a gene encoding an enzyme for the transport or processing of a polysaccharide or oligosaccharide unit,  
30     the gene being involved in the synthesis of a *E. coli* O antigen, in conditions sufficient to allow the pair of nucleic acid molecules to hybridise to the gene;

- (b) contacting a gene of an *E. coli* in the sample with a pair of nucleic acid molecules according to claim 1  
35     in conditions sufficient to allow the pair of nucleic acid molecules to hybridise to the gene; and

- (c) detecting pairs of nucleic acid molecules which

- 86 -

are hybridised to the genes, to detect the H and O serotype of the *E. coli* in the sample.

16. A method according to claim 15 wherein the pair of nucleic acid molecules of step (a) is selected from the group consisting of:

*wbdH* (nucleotide position 739 to 1932 of Figure 5),  
*wzx* (nucleotide position 8646 to 9911 of Figure 5),  
*wzy* (nucleotide position 9901 to 10953 of Figure 5),  
10 *wbdM* (nucleotide position 11821 to 12945 of Figure 5),  
*wbdN* (nucleotide position 79 to 861 of Figure 6),  
*wbdO* (nucleotide position 2011 to 2757 of Figure 6),  
*wbdP* (nucleotide position 5257 to 6471 of Figure 6),  
*wbdR* (nucleotide position 13156 to 13821 of Figure 6),  
15 *wzx* (nucleotide position 2744 to 4135 of Figure 6) and  
*wzy* (nucleotide position 858 to 2042 of Figure 6).

17. A method according to claim 15 wherein the nucleic acid molecules of the pair of step (a) are primers  
20 selected from the group of primers shown in Tables 8, 8A, 9 and 9A.

18. A method according to claim 15 wherein the hybridised pairs of nucleic acid molecules are detected  
25 by the polymerase chain reaction.

19. A method for detecting the H and O serotype of *E. coli* in a sample, the method comprising the following steps:

30 (a) contacting a gene of an *E. coli* in the sample with a nucleic acid molecule according to claim 1, in conditions sufficient to allow the nucleic acid molecule to hybridise to the gene; and

(b) detecting a nucleic acid molecule which is  
35 hybridised to the gene, to detect the H and O serotype of *E. coli* in the sample.

- 87 -

20. A method according to claim 19 wherein the nucleic acid molecule is according to any one of SEQ ID NOS: 9, 55, 57 to 65.

5 21. A method according to any one of claims 8, 9, 11, 15 or 19 wherein the sample is selected from the group consisting of a sample derived from food, a sample derived from faeces and a sample derived from a patient or animal.

10 22. A kit for identifying the H serotype of *E. coli*, the kit comprising at least one nucleic acid molecule according to any one of claims 1 to 6.

15 23. A kit for identifying the H and O serotype of *E. coli*, the kit comprising:

(a) at least one nucleic acid molecule according to any one of claims 1 to 6; and

20 (b) at least one nucleic acid molecule derived from and specific for a gene encoding a transferase or a gene encoding an enzyme for the transport or processing of a polysaccharide or oligosaccharide unit, the gene being involved in the synthesis of a particular *E. coli* O antigen.

25 24. A kit according to claim 23 wherein the at least one nucleic acid molecule of (a) is selected from the group consisting of:

*wbdH* (nucleotide position 739 to 1932 of Figure 5),

*wzx* (nucleotide position 8646 to 9911 of Figure 5),

30 *wzy* (nucleotide position 9901 to 10953 of Figure 5),

*wbdM* (nucleotide position 11821 to 12945 of Figure 5),

*wbdN* (nucleotide position 79 to 861 of Figure 6),

*wbdO* (nucleotide position 2011 to 2757 of Figure 6),

*wbdP* (nucleotide position 5257 to 6471 of Figure 6),

35 *wbdR* (nucleotide position 13156 to 13821 of Figure 6),

*wzx* (nucleotide position 2744 to 4135 of Figure 6) and

*wzy* (nucleotide position 858 to 2042 of Figure 6).

- 88 -

25. A kit according to claim 24 wherein the nucleic acid molecule of (a) is a primer selected from the group of primers shown in Tables 8, 8A, 9 and 9A.

5

10  
15  
20  
25  
30  
35  
40  
45  
50  
55  
60  
65  
70  
75  
80  
85  
90  
95  
100  
105  
110  
115  
120  
125  
130  
135  
140  
145  
150  
155  
160  
165  
170  
175  
180  
185  
190  
195  
200

1/96

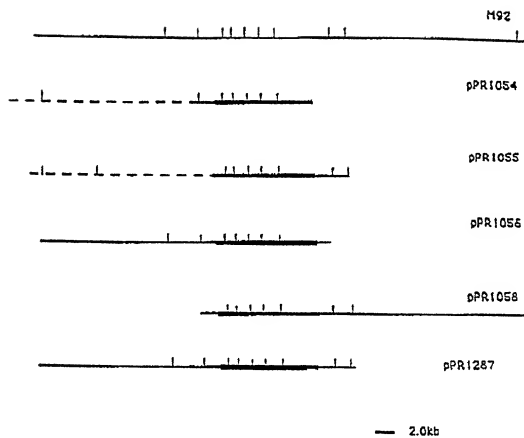
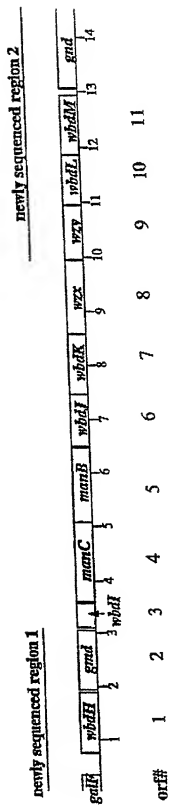


Figure 1



Figure 2



### Figure 3

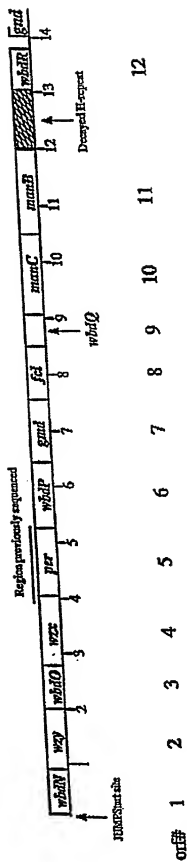


Figure 4

GATCTGATGGCCGTAGGGCGCTACGTGCTTCTGCTGATATCTGGGCTGAGTTGGAAAAA	60
ACTGCTCCAGGTGCCTGGGGACGTATTCAACTGACTGATGCTATTGCAGAGTTGGCTAAA	120
AAACAGTCTGTGATGCCATGCTGATGACCGGCGACAGCTACGACTGCGGTAAGAAGATG	180
GGCTATATGCAGGCATTTCGTTAAGTATGGGCTGCGCAACCTTAAAGAAGGGGCGAAGTTC	240
CGTAAGAGCATCAAGAAGCTACTGAGTGAGTAGAGATTTACACGCTTTTGTGACGATAAG	300
CCAGAAAAATAGCGGCAGTTAATCATCCAGGCTTCTATGCTTTAAGCAATGGAATGTTAC	360
TGCCGTTTTTTTATGAAAAATGACCAATAATAACAAGTTAACTACCAAGTTTAATCTGCT	420
TTTGTGTTGGATTTTCTTGTGTTCTGGTGCATTTGGTAAGACAATTAGCGTGAGTTTTTA	480
GAGAGTTTTCGGGATCTCGCGGAATGCTACATCTTTGGCATTAGTTAGTGCACTGG	540
TAGCTGTTAAGCCAGGGGCGGTAGCTTGCTTAATTAATTTTAAAGTATACATTTATTCT	600
TGCCGCTTATAGCAATAAAGTCAATCGGATTAACCTCTTTTCCATTAGGTAAAAGAGT	660
GTTGTAGTCGCTCAGGGAATTTGGTTTTGGTAGTAGTACTTTTCAAATTATCCATTTTC	720
<p style="text-align: center;">Start of orf1 M L L C C I H I N V Y Y L L</p>	
CGATTTAGATGGCAGTTGATGTTACTATGCTGCATACATATCAATGTATATTATTACTT	780
L E C D M K K I V I I G N V A S M M L R	
TTAGAATGTGATATGAAAAAATAGTGATCATAGGCAATGTAGCGCTCAATGATGTTAAGG	840
F R K E L I M N L V R Q G D N V Y C L A	
TTCAGGAAAGAAATTAATCATGAATTTAGTGAGGCAAGTGATAATGTATATTGCTGACGA	900
N D F S T E D L K V L S S W G V K G V K	
AATGATTTTCCACTGAAGATCTTAAAGTACTTTCGTCATGGGCGTTAAGGGGGTTAAA	960
F S L N S K G I N P F K D I I A V Y E L	
TTCTCTCTTAAGTCAAAGGGTATTAATCCTTTTAAAGGATATAATTTGCTGTTATGAAGTCTA	1020
K K I L K D I S P D I V F S Y F V K P V	
AAAAAATTTCTTAAGGATATTCCCCAGATATTGTATTTTCATATTTTGTAAAGCCAGTA	1080
I F G T I A S K L S K V P R I V G M I E	
ATATTGGAACATTGCTTCAAAGTTGTCAAAGTGCCAAAGGATTGTTGGAATGATTGAA	1140
G L G N A F T Y Y K G K Q T T K T K M I	
GGTCTAGTAAATGCTTCACTTATTATTAAGGAAAGCAGACCAAAAACTAAAATGATA	1200
K W I Q I L L Y K L A L P M L D D L I L	
AAGTGGATACAAATTTCTTTATATAGTTAGCATTACCGATGCTGTAGTATTGATTCTA	1260
L N H D D K K D L I D Q Y N I K A K V T	
TTAATCATGTAGTATAAAAAGATTAACTCGATCAGTATAATATAAGCTAAGGTAAACA	1320
V L G G I G L D L N E F S Y K E P P K E	
GTGTTAGTGGGATTTGGATCTTAATGAGTTTTTCATATAAAGAGCCACCAGAAAGG	1380
K I T F I F I A R L L R E K G I F E F I	
AAAAATACCTTTATTTTATAGCAAGGTTATTAAGAGAGAAAGGATATTGAGTTTATT	1440
E A A K F V K T T Y P S S E F V I L G G	
GAAGCCGCAAGTTTCGTTAAGACAACCTATCCAAGTTCTGAATTTGTAATTTAGGAGGT	1500

WO 99/61458

PCT/AU99/00385

6/96

F E S N N P F S L Q K N E I E S L R K E 1560  
 TTGAGAGTAATAATCCTTTCTCATTACAAAAAATGAAATGGAATCGCTAAGAAAAGAA  
 H D L I Y P G H V E N V Q D W L E K S S 1620  
 CATGATCTTATTATCCTGGTCATGTGGAAAAATGTTCAAGATTGGTTAGAGAAAAGTTCT  
 V F V L P T S Y R E G V P R V I Q E A M 1680  
 GTTTTTGTTTACCTACATCATATCGAGAAGGCGTACCAAGGGTGATCCGAAGAAGCTATG  
 A I G R P V I T T N V P G C R D I I N D 1740  
 GCTATTGCTAGACCTGTAATAACAACATAATGTACCTGGGTGTAGGGATATAATAAATGAT  
 G V N G F L I P P F E I N L L A E K M K 1800  
 GGGGCAATGGCTTTTGTACCTCCATTGAAATTAATTACTGGCAGAAAAAATGAAA  
 Y F I E N K D K V L E M G L A G R K F A 1860  
 TATTTTATTGAGAAATAAGATAAAGTACTCGAAATGGGGCTTGCTGGAAGGAAGTTGCA  
 E K N F D A F E K N N R L A S I I K S N 1920  
 GAAAAAACTTTGATGCTTTTGAAAAAATAATAGACTAGCATCAATAATAAATCAAAAT

## End of orf1

N D F \*

AATGATTTTGACTTGAGCAGAAAAATATTTATATTTCAATCTGAAAAATAAGGCTGTTA 1980

## Start of orf2

M N K V A L I T G I T G O D G S Y L A 2040  
 TTATGAATAAAGTGGCATTATTAATCTGATCACTGGGCAAGATGGCTCCTATTGGCAG  
 E L L L E K G Y E V H G I K R R A S S F 2100  
 AATTATGTTAGAAAAAGGTTATGAAGTTCATGGTATTAACGCCGTGCATCTTCATTTA  
 N T E R V D H I Y Q D S H L A N P K L F 2160  
 ATACTGAGCGAGTGGATCACATCTATCAGGATTCACATTTAGCTAATCTCAAACTTTTC  
 L H Y G D L T D T S N L T R I L K E V Q 2220  
 TACACTATGGCGATTTGACAGACTTCCAATCTGACCCGTATTTAAAAAGAGTTCAAC  
 P D E V Y N L G A M S H V A V S F E S P 2280  
 CAGATGAAGTTTACAATTTGGGGCGATGAGCCATGTAGCGGTATCATTTGAGTCACCAG  
 E Y T A D V D A I G T L R L L E A I R I 2340  
 AATACACTGCTGATGTTGATGCGATAGGAACATGCGCTCTCTTGAAGCTATCAGGATAT  
 L G L E K K K T K F Y Q A S T S E L Y G L 2400  
 TGGGCTGGAAAAAAGACAAAAATTTATCAGGCTCAACCTTCAGAGCTTTATGGTTTGG  
 V Q E I P Q K E T T P F Y P R S P Y A V 2460  
 TTCAGAAATTCACAAAAAGAGACTACGCCATTTTATCCACGTTCGCCCTTATGCTGTG  
 A K L Y A Y W I T V N Y R E S Y G M F A 2520  
 CAAATTATATGCGTATTTGGATCACGTGTAATTAATCTGAGTCTTATGGTATGTTGCTT  
 C N G I L F N H E S P R R G E T F V T R 2580  
 GCAATGGTATCTCTTTAACCACGAATCACCTCGCCGTGGCGAGACCTTTGTACTCGTA  
 K I T R G I A N I A Q G L D K C L Y L G 2640  
 AAATAACACGGGGATAGCAAATATTGCTCAAGGCTTGATAAATGCTTACTTGGGAA  
 N M D S L R D W G H A K D Y V K M Q W M 2700  
 ATATGGATTCTCTCGGTGATGGGACATGCTAAGGATTATGTCAAAATGCAATGGATGA

WO 99/61458

PCT/AU99/00385

7/96

M L Q Q E T P E D F V I A T G I O Y S V 2760  
 TGCTGCAGCAAGAACTCCAGAAGATTTTGTAATTGCTACAGGAATTCATATATTCTGTCC  
 R E F V T M A A E Q V G I E L A F E G E 2820  
 GTGAGTTTGTCCAAATGGCGCAGAGCAAGTAGGCATAGAGTTAGCATTTGAAGGTGAGG  
 G V N E K G V V V S V N G T D A K A V N 2880  
 GAGTAAATGAAAAGGTGTGTGTGTTTCGGTCAATGGCACTGATGCTAAAGCTGTAAACC  
 P G D V I I S V D P R Y F R P A E V E T 2940  
 CGGGCGATGTAATTATATCTGTAGATCCAAGGATTTTAGGCCCTGCAGAGTTGAAACCT  
 L L G D P T N A H K K L G W S P E I T L 3000  
 TGCTTGGCGATCCTACTAATGCGCATAAAAAATTAGGATGGAGCCCTGAAATTACATTGC  
 R E M V K E M V S S D L A I A K K N V L 3060  
 GTGAAATGGTAAAAGAAATCGPTTCCAGCGATTAGCAATAGCGAAAAAGACCTCTTCG  
 End of orf2  
 L K A N N I A T N I P Q E \*  
 TGAAGCTAATAACATTGGCACTAATATTCCGCAAGAAATAAAAAGATAATACATTAAAT 3120  
 Start of orf3  
 M F  
 AATTAATAATGGTGTAGATTTATTAGTACCATTATTTTTTTTTTGGGTGACATAATGTTTA 3180  
 I T S D K F R E I I K L V P L V S I D L 3240  
 TACATAGCTAAATTTTACAGAAATATCAAGTTAGTTCCTTACATATCAATTTGATCTGC  
 L I E N E N G E Y L F G L R N N R P A K 3300  
 TANTTGAACCGAGATGGTGAATATTTATTTGGTCTTAGGAATATTCGACCGCGCAAAA  
 N Y F F V P G G R I R K N E S I K N A F 3360  
 ATTATTTTTTTGGTTCGAGGTGGTAGGATTCGCAAAAAATGAATCTATTAATAATGCTTTTA  
 K R I S S M E L G K E Y G I S G S V F N 3420  
 AAGAATATCATCTATGGAATTAGGTAAAGAGTATGGTATTTTCAGGAAGTGTPTTTAATG  
 G V W E H F Y D D G F F S E G E A T H Y 3480  
 GTGTATGCGAACATTTCTATGATGATGGTTTTTTTTTTCTGAAGGCGAGCGAACACATTATA  
 I V L C Y T L K V L K S E L N L P D D Q 3540  
 TAGTGETTTTGTACAGACTGAAGTTCTTAAAAAGTAATGAAATCTCCAGATCTCAAC  
 H R E Y L W L T K H Q I N A K Q D V H N 3600  
 ATCTGTAATACCTTTGGCTAACTAAACCGCAATAAATGCTAAAGAGATGTTCTATACCT  
 End of orf3  
 Start of orf4  
 M  
 Y S K N Y F L \*  
 ATTCAAAAATATATTTTGTGTAATTTTATATAAAAAATTAATATCGCAGAGAAATTTGATGTT 3660  
 S Q C L Y P V I I A G G T G S R L W P L 3720  
 CTCATATGCTTTTACCGCTGTAAATATTCGCGGAGGAACCGGAGCGCTATATGCGCGTTGT  
 S R V L Y P K Q F L N L V G D S T M L Q 3780  
 CTCGAGTATATACCTTAACAAATTTTAAATTTAGTTGGGGAATCTCAAAATTTGCAAA  
 T T I T R L D G I E C E N P I V I C N E 3840  
 CAAATATACCGCTTTGGATGCGATCGAATGCGAAATCGCAATGTTATCTCGAATGAAG  
 D H R F I V A E Q L R Q I G K L T K N I 3900  
 ATCAACCGATTTATATGTCAGAGAGCAATACGACAGATTGCTAAGCTAACCAAGAAATATA  
 I L E P K G R N T A P A I A L A A F I A 3960  
 TACTTGAAGCGCAAGGCGGTAAATCTGCACTGCGCATAGCTTTAGCTGCTTTTATCGCTTC

8/96

Q K N N P N D D P L L L V L A A D H S I 4020  
 AGAAGAAATAATCTAATGACGACCCCTTATATATAGTACTGCGGCAGACCACTCTATAA  
 N N E K A F R E S I I K A M P Y A T S G 4080  
 ATAAATGAAAGCATTTCGAGAGTCAATAATAAAGCTATGCGGTATGCAACTTCTGGGA  
 K L V T F G I I P D T A N T G Y G Y I K 4140  
 AGTTGATCAACTTGGGAATTTATTCGGGACACGGCAATACTGGTTATGGATATATTAAGA  
 R S S S A D P N K E F P A Y N V A E F V 4200  
 GAAGTTCCTCAGCTGATCCTAATAAAGAAATTCGCAGCATATAATGTTGCGGAGTTTGTAG  
 E K P D V K T A Q E Y I S S G N Y Y W N 4260  
 AAAACCCAGATGTTAAAAACAGCACAGGAATATATTTTCGAGTGGGAATTTATTAACGGAAAT  
 S G M F L F R A S K Y L D E L R K F R P 4320  
 GCGGAATGTTTTTATTTTCGCGCCAGTAAATATCTTGATGAACACGGAATTTAGACCCAG  
 D I Y H S C E C A T A T A N I D M D F V 4380  
 ATATTTATCATAGCTGTGAATGTCGCAACCGCTACAGCAAAATATAGATATGGAATTTGTGTC  
 R I N E A E F I N C P E E S I D Y A V M 4440  
 GAATTAACGAGCGGTGAGTTTATTAATTTGTCCTGAAGACTCTATTCGATTATGCTGTGATGG  
 E K T K D A V V L P I D I G W N D V G S 4500  
 AAAAAACAAAGAGCGGTGATGTTTCTTCGGATAGATATTTGGCTGGGAATGACCGTGGGTCTCT  
 W S S L W D I S Q K D C H G N V C H G D 4560  
 GGTCTACACTTTGGGATATAAGCCAAAGGATTTGCCATGTTAATGCTGTGCCATGGGGATG  
 V L N H D G E N S F I Y S E S S L V A T 4620  
 TGCTCAATCATGATGGAGAAATAGCTTTTATTTACTCTGAGTCAAGTCTGGTTGCGACAG  
 V G V S N L V I V O T K D A V L V A D R 4680  
 TCGGAGTAAGTAATTTAGTAAATGTCGCAACCAAGGATGCTGTACTGGTTGCGGACCGGTG  
 D K V Q N V K N I V D D L K K R K R A E 4740  
 ATAAAGTCCAAAAATGTTAAAAACATAGTTGACCATCTAAAAAGAGAAACAGTGTGTAAAT  
 Y Y M H R A V F R P W G K F D A I D Q G 4800  
 ACTACATGCATCGTGGAGTTTTTGGGCGCTTGGGGTAAATTCGATGCAATAGACCAAGGGG  
 D R Y R V K K I I V K P G E G L D L R M 4860  
 ATAGATATAGAGTAATAAAAAATATAGTTAAACCAAGGAGAGGGTTAGATTTAAGGATGC  
 H H H R A E H W I V V S G T A K V S L G 4920  
 ATCATCATAGCCAGCAGCATTCGATTGTTGTAATCCCGTACTGCTAAAGTTTCACTAGGTA  
 S E V K L L V S N E S I V I P Q G A K Y 4980  
 GTGAAGTTAAACTATAGTTTCTAATGAGTCTATATATATCCCTCAGGGAGCAAAATATA  
 S L E N P G V I P L H L I E V S S G D Y 5040  
 GTCTTGAGAAATCCAGGCGTAATACCTTTGCATCTAATTAAGTAAGTCTGCTGATTAACE  
 L E S D D I V R F T D R Y N S K Q F L K 5100  
 TGAATTCAGATGATATAGTGGCTTTTACTGACAGAGATATACAGTAACAAATTCCTAAAGC

End of orf4 Start of orf5

R D \* M N K I T C F K A Y D I R G R L

GAGATTCATATAATATGAATAAATAACTTGCCTTCAAGCATATGATATACCTGGGCGCTCT 5160

G A E L N D E I A Y R I G R A Y G E F F 5220  
 TGGTGTGAATGAATGATGAAATAGCATATAGAAATGGTGGCGCTTATGGTGAATTTTT  
 K P Q T V V V G G D A R L T S E S L K K 5280  
 TAAACCTCAAACTGTAGTTGTGGGAGGAGATGCTGGCTTAACAAGTGAGAGTTTAAAGAA  
 S L S N G L C D A G V N V L D L G M C G 5340  
 ATCATTCTCAAAATGGGCTATGTGATGCGAGGCGTAAATGTCTTGTAGATCTTGGAACTGTGG  
 T E E I Y F S T W Y L G I D G G I E V T 5400  
 TACTGAAGAGATATATTTTCCACTTGGTATTAGGAATTGATGGTGGAAATCGAGCTAAC  
 A S H N P I D Y N G M K L V T K G A R P 5460  
 TCGAAGCCATAATCCAAATTCATTATATATGGAAATGAATTAGTAACCAAAAGGTGCTCGAAC  
 I S S D T G L K D I Q Q L V E S N N F E 5520  
 AATCAGCAGTGACACAGGTCTCAAAAGATATACAAACAATTACTAGAGAGTAATAATTTTGA  
 E L N L E K K G N I T K Y S T R D A Y I 5580  
 AGAGCTCAACCTAGAAAAAAAAGGGAATATTACCAAAATATTCCACCCGAGATGCTTACAA  
 N H L M G Y A N L Q K I K K I K I V V N 5640  
 AATATCTTGTAGTGGCTATGCTAACTGCGAAAAAATAAAAAATCAAAATGATTTGTGAA  
 S G N G A A G P V I D A I E E C F L R N 5700  
 TTCTGGGAATGGTGGAGCTGGCTGCTGTTATGTATGCTATTGAGGAATGCTTTTACCGAA  
 N I P I Q F V K I N N T P D G N F P H G 5760  
 CAATATTCCGATTCAGTTTGTAAAAATAAATAATACACCCGATGGTAATTTTCCACATGG  
 I P N P L L P E C R E D T S S A V I R H 5820  
 TATCCCTAATCCATTACTACCTGAGTGCAGAGAGATACACAGTCCGTTTATAAGACA  
 S A D F G I A F D G D F D R C F F D E 5880  
 TAGTGCTGATTTTGGTATTGGATTGATGGTGGATTTTTGAATGGTGTTTTTTCTTTGATGA  
 N G Q F I E G Y Y I V G L L A E V F L G 5940  
 AATATGGCAATTTATTGAAAGGATACACATGCTTGGTTTATTAGCGGAAGTTTTTTAGG  
 K Y P N A K I I H D P R L I W N T I D I 6000  
 GAAATATCCAAAGCAAAATTCATTTCATGATTCCTGGCGTTATATGGAATACATTATGATAT  
 V E S H G G I P I M T K T G H A Y I K Q 6060  
 CGTAGAAATGCTGGTGGTATACCTATAATGACTAAAACCGGCTCATGCTTACATTAAGCA  
 R M R E E D A V Y G G E M S A H H Y F K 6120  
 AAGATTCCTGGAAGAGGATGCCGTATATGGCGCGCAATGAGTGGCGCATCATTTATTA  
 D F A Y C D S G M I P W I L I C E L L S 6180  
 AGATTTTGGATACTGGGATAGTGGAAATGATTCCTTGGATTTTAATTTGTGAACCTTTTGA  
 L T N K K L G E L V C G C I N D W P A S 6240  
 TCTGCAAAATAAAAAATTAGGTGAACCTGGTTGTGGTTGTATAAACGACTGGCGCGCAAG  
 G E I N C T L D N P Q N E I D K L F N R 6300  
 TGGAGAAATAAAGCTGACACTAGACAATCCGCAAAATGAATAAGATATAAATATTATTAAG  
 Y K D S A L A V D Y T D G L T M E F S D 6360  
 TTAACAAAGATAGTGGCTTAGCTGTGTGATTACACTGATGGATTAACATATGGAGTTCTCTGA  
 W R F N V R C S N T E P V V R L N V E S 6420  
 TTGGCGTTTTAATGTAGATGCTCAAAATACAGAACTGTAGTACGATGAATGTGAATC  
 R N N A I L M O E K T E E I L N F I S K 6480  
 TAGGAATAATGCTATTCTTATGCGAGGAAAAACAGAGAAATTTGTGAATTTTATATCAAA



10/96

End of orf5	Start of orf6	
	M K V L L T G	
ATAAATTTCGACCTGAGTTCATAATGGGAACAAGAAATATATGAAAGTACTTCTGACTGG		6540
STG M V G K N I L E H D S A S K Y N I		
CTTAAGCTGGCATGCTTGGTAAGAAATATATTAGAGCATGATAGTGCAGATAAATATATAAT		6600
L T P T S S D L N L L D K N E I E K F M		
ACTTACTCCAACAGCTCTGTATTGAATTTATATAGATAAAATGAAATAGAAAAATTCAT		6660
L I N M P D C I I H A A G L V G G I H A		
GCTATCTCAACATGCCAGACTGTATTTATACATGCAGCGGGATTAAGTTGGAGGCCATTCATGC		6720
N I S R P F D F L E K N L Q M G L N L V		
AAATATAAGCAGGCGCTTTGATTTTCTGGAAAAAAATTTTCAGATGGGTTTAAATTTAGT		6780
S V A K K L G I K K V L N L G S S C M Y		
TTCCCTCGCAAAAAAAGGTATCAAGAAAAGTGGCTTAAGTTGGGTAGTTTCATGCATGTA		6840
P K N F E E A I P E K A L L T G E L E E		
CCCCAAAAAAGCTTTGAAGAGGCTATTCCTGAGAAAAGCTCTGTTAACTGGTGAGCTAGAGA		6900
T N E G Y A I A K I A V A K A C E Y I S		
AACTAATGAGGGATATGCTATTGCGAAAAATTCGTGTAGCAAAAAGCATGCGAATATATATC		6960
R E N S N Y F Y K T I I P C N L Y G K Y		
AAAGAGAAAACTCTAATTTATTTTATAAAACAAATTAACCCCATGTAAATTTATATGSGAAAA		7020
D K F D D N S S H M I P A V I K K I H H		
TGATAAATTTGATGATACTCGTCACATATGATTCGCGGAGTTATAAAAAAATCCATCA		7080
A K I N N V P E I E I W G D G N S R R E		
TGGCAAAATTAATAATGTCCGAGAGATCGAAATTTGGGGGGGATGGTAATTCGCGCGCTGA		7140
F M Y A E D L A D L I F Y V I P K I E F		
GTTTATGTATGCAAGATTTTACGTGATCTTATTTTATGCTTATTCCTAAAAATAGAAAT		7200
M P N M V N A G L G Y D Y S I N D Y Y K		
CATCGCTAATATGCTAAATGCTGCTTAGGTTAGCATATTCGAATTAATGACTATTATAA		7260
I I A E E I G Y T G S F S H D L T K P T		
GATTAATTCGAGAGAAATTCGTATATCTGGGAGTTTCTCTCATGATTTAAACAAAAGCAAC		7320
G M K R K L V D I S L L N K I G W S S H		
AGGAATGAACCGAAGCTAGTAGATATTTCTATTCCTTAATAAAATTCGCTTGGTCAAGTCA		7380
F E L R D G I R K T Y N Y Y L E N Q N K		
CTTTGAAGCTCAGAGATGGCATCAGAAAAGACATATAATTTATTTACTTGGAGAAATCAAAATA		7440

Start of orf7, End of orf6

M I T Y P L A S N T W D E Y E Y A A I Q	
ATGATACATACCCACTTGTGTATTAATCTTGGGATGAATATGAGTATGCAGCAATACAG	7500
S V I D S K M F T M G K K V E L Y E K N	
TCAGTAATTAAGCTCAAAAAATGTTTACCATGGGTAAAAAGGCTTGAGTTATATAGAAAAAT	7560
F A D L F G S K Y A V M V S S G S T A N	
TTTGCTGATTTGTTTGGTAGCAAAATATGCGGTAAATGGTTAGCTCTGCTTCTACAGCTAA	7620

11/96

TGTTTAAATGAAGATTGGCTGCCCTTTTCCTCACTAATAAACCAGAAACCTTAAGAAGAGGTGTATGA	
I I V P A V S W S T T Y Y P L Q Q Y V G L	7680
ATAATAGTAGTACCTGCAGCTGTGCATGGCTACGACATATTACCCTCTGCAACAGATATGGCCTTA	7740
K V K F V D I N K E T L N I D S L K	
AAGGTGAAGCTTTGTGCGATATCAAATAAGAAACTTTAATATTGCATATCGATAGTTTGAAA	7800
N A I S D K T K A I L T V N N L G N P N	
AATGCTATTTCAGATATAAACAAAGCAANTATTGCACSTAAATTTATTAGSGTAATCCCTAAT	7860
D F A K I N E I N N R D I I L L E D N	
GATTTTGGCAAAAAATATGAGATAATAAATAATAGGGATATTACTTTACTTAGAAGCATAC	7920
C E S M G A V F O N K Q A G T T F G V M G	
TGTGAGTGCATGGGCGCGGCTCTTTCAAAAATAAGCAGCGCAGGCACATTCGGAGTTATGGGT	7980
T F S S F Y S H H I A T M E G G C V U T	
ACCTTTAGTTCTTTTACTCTCATCATATAGTACAATGGAAAGGGGGCTGGCTAGTCTACT	8040
D D E E L Y H V L L C L R A H G W T R N	
GATGATGAAAGAGCTGTATCATGTATTTGTTCTGCGCTTCGAGGCTCATGGTTGGCAAGAAAT	8100
L P K E N M T G V T G T K S D D I F E E S F	
TTACCAAAAGAGAAATATGGTTTACAGGCACTAAGAGTGATGATATTTTTCGAAGAGTCTGTT	8160
K F V V L P G Y A N V R P L E M S G A I G I	
AAGTTTGTGTTTTACCCAGGATACAATGTTTCGCCCACTTGAATAGAGTGGTGCTATTGGGATA	8220
E Q L K K L P G F I S T R R S N A Q Y F F	
GAGCAACTTTAAAGTTTACCAGGTTTTATATTCGCCACAGAGCTTCCAATGCGACAATATTTT	8280
V D K F K D H P F L D I Q K E V G E S S	
GTACATAAATTTAAGATCACCCATTCCTTGATATACAAAAGAGTTGGTGAAAGTAG	8340
W F G F S F V I K E G A A I E R K S L V	
TGTTTGTGTTTTTCCCTTCGTTATAAAGCAGGGAGCTGCTATTGACAGGAAAGGTTTAGTA	8400
N N T L I S A G E I E C R P I V T G T G N F L K	
AATPATECTGATCGACAGCGATTGCAATCCGACCPAATTTGTTACTGGGAATTTTCTCAA	8460
N E R V L S Y F D Y S V H D T V A N A E	
AATGAACGTTGTTTGAAGTTATTTTGATTTACTCTGTGATCATGATAGCGTATGCAATATGCCGA	8520
Y I A D K N G F F V G N H Q I P L F N E I	
TATATGATTAAGAATGGTTTTTTCGCGAAACACAGGATACCTTTGTTTATFNAATA	8580

End of orf7

D Y L R K V L K \*  
 GATTATCTACGAAAGTATTAAAACTACTACGAGGCACTCTATTTCGAATAGAGTCTCT 8640

Start of orf8

M V L P V K K I L P F G Y S K V L P  
 TTAAGATCGTATTACAGCTGAAATAAATTTAGCGTTTGGCTATTCTTAAGTATATACCAAC 8700  
 P V I E Q F V N P I C I F I I T P L I L  
 CGGTATTATGAACAGTTTGTCATATCAATTTGGCATCTTCATATATACACCACTAATATCTCA 8760  
 N H L G K Q C S Y G N W I L L I T I V S F  
 ACCAATCGGGTACGCAAAAGCTATGCTATATGGATTTTATATATAGTATATATATCTCTTTT 8820

12/96

S Q L I C G G C S A W I A K I I A E Q R  
 CTCAGTTAATAATGTCGGAGGATGTTCCGCGATGGATTGCAAAAAATCATTGCGAGACAGAGAA 8880  
 I L S D L S K K N A L R Q I S Y N F S I  
 TTCTTAGTGATTATCAAAAAAATGCTTTACGCTCAATTTCTATATAATTTTTCGAATTG 8940  
 V I I A F A V L I S F L I L S I C F F D  
 TTATATATCCGCAATTTGCGGTAATGATTGCTTTTCTTAATATAAGTATTTGTTCTGCTAGT 9000  
 V A R N N S S F L F A I I I C G F F Q E  
 TTCGGAGGAATAATTTCTCATTCTTATTGCGGATTATTTATTTGTGGTTTTTTTCAGGAAG 9060  
 V D N L F S G A L K G F E K F N V S C F  
 TTGATAATTTTATTTAGTGGTGGCTTAAAGGTTTTGAAAAATTAATGTAATCATGTTTTT 9120  
 F E V I T R V L W A S I V I Y G I Y G N  
 TTGAAGTAATTTACAGAGTGGCTCTCGGCTTCATAGTAATAATGCGCATTTCGGAATG 9180  
 A L L Y F T C L A F T I K G M L K Y I L  
 CACTGTATATTTTACATGTTTAGCGCTTTACCATTAAGGTAATGCTAAAAATATATTTG 9240  
 V C L N I T G C F I N P N F N R V G I V  
 TATGTCGCAATATTTACCGGTGTTTTCATCAATGCTAATTTTAAATAGAGTTGGGATTGTTA 9300  
 N L L N E S K W M F L Q L T G G V S L S  
 ATTTGTAAATGAGTCAAAATGGATGTTTCTTCGAATTAACGTGGTGGGCTCTCACTTAGTT 9360  
 L F D R L V I P L I L S V S K L A S Y V  
 TGTGTGATAGGCTCGTAATACCATTTGATTTTATCTGTCAAGTAACGTGGCTTTATGTCG 9420  
 P C L Q L A Q L M F T L S A S A N Q I L  
 CTTGCGCTTCACTCAATTCATGTTCACTCTTCTGCGCTGCAAAATCAAAATATTAC 9480  
 L P M F A R M K A S N T F P S N C F F K  
 TAGCAATGTTTGTAGATAAAGCATCTTAACACATTTGCGCTCAATTTGTTTTTTTAAAA 9540  
 I L L V S L I S V L P C L A L F F F G R  
 TTCTGCTGTATCACTAATTTCTGTTTTGCGCTGTCTTGGCTTATCTTTTTTGGTGGTG 9600  
 D I L S I W I N P T F A T E N Y K L M Q  
 ATATATTTATCAATATGGATAAAGCTAGATTTCGAACGAAATATATAAATTAATGCAAA 9660  
 I L A I S Y I L L S M M T S F H F L L L  
 TTTTAGCTATAAGTTACATTTTATTGTCAATGATGACATCTTTTCATTCTGTTGTTAG 9720  
 G I G K S K L V A N L N L V A G L A L A  
 GAATTCGTAATCTAAGCTTCTTGCAAATTTAAATCTGCTTGCAGGGCTGCGACTTGTG 9780  
 A S T L I A A H Y G L Y A I S M V K I I  
 CTTCAAGCTTATTCGAGCTCATTTATGGCCTTTATGCAATATCTATGCTAAAAATTAAT 9840  
 Y P A F Q F Y Y L Y V A F V Y F N R A K  
 ATCCGGCTTTTCAATTTATAGCTTTATGTAGCTTTTGTCTATTTTAAATAGCGGAAAA 9900

Start of orf9, End of orf8

M S I D L L F S I T E I A I V F S C T I  
 N V Y  
 ATGCTCATTTGATTACTTTTTTCAATTAAGTGAATTCGCAATTTGTTTTCTTGCACATT 9960  
 Y I F T C Q L L M R R I Y L D K S I L I  
 TACATATTTACTCAATGTTTGTAAATGCGGAGGATCTATTTAGATAAAAGTATTTTAATT 10020  
 L L C L L F F L V I I O L P E L N V N G  
 CTTTATGCTTGCTCTTTTTTTTAGTAATCATTCACCTTCTGAGCTTAATGTAAACGGT 10080

13/96

L V D S L K L S L P L L M V F I A F Q K TTGGTCGATTCTTTAAAGTTATCACTGCCTTTATTGATGGCTCTTTATCGCTTTTCAAAA	10140
P K L C L W V I I A L L F L N S A F N F CCGAAATTATGCTTGTGGGTTATTATTGCATTGTTGTTTTGAACCTCGCATTTAATTTT	10200
L Y L K T F D K F S S F P F T F F I L L TTATATTTAAAGACATTCGATAAGTTTAGCTCATTTCCTTTTACTTTTTTATATTGCTG	10260
F Y L F R L G I G N L P V Y K N K K F Y TTTACTTGTTTAGATTGGGAATTGGTAATTTACCGTTTATAAAAATAAAAATTTTAC	10320
A L I F L F I L I D I M Q S L L I N Y R CGCTTGATTTTCTCTTTATATTAATAGACATAATGCAGTCATTGTTAATAAATTATAGG	10380
G Q I L Y S V I C I L I L V F K V N L R GGGAGATTTTATATCCGTAATTTGCATCTGATACTTGTGTTTAAAGTTAATTTAAGA	10440
K K I P Y F F L M L P V L Y V I I M A Y AAAAAGATTCCATACCTTTTTTAAATGCTGCCAGTTTATATGTAATTTATATGGCTTAT	10500
I G F N Y F N K G V T F F E P T A S N I ATTGGTTTAAATTATTCAATAAAGGCGTAACCTTTTTTGAACCTACAGCAAGTAATATT	10560
E R T G M I Y Y L V S Q L G D Y I F H G GAACGTACGGGATGATATATTATTGGTTTTCACAGCTTGGTGATATATATTCCATGGT	10620
M G T L N F L N N G G Q Y K T L Y G L P ATGGGACATTAAATTTCTTAAATAACGGCGGACAATAAAGACCTTATATGGACCTTCCA	10680
S L I P N D P H D F L L R F F I S I G V TCATTAATTCCTAATGACCCATGATTTTTTATTACGGTTCTTTATAAGTATGGTG	10740
I G A L V Y H S I F F V F F R R I S F L ATAGGAGCATTTGTTTATCATCTATATTTTTTGTTTTTTATAGGAGAATATCTTTCTTA	10800
L Y E R N A P F I V V S C L L L L Q V V TTATATGAGAGAAATGCTCCTTCATTTGTTGTAAGTTGTTGTTTACGTGTACAAGTTGTG	10860
L I Y T L N P F D A F N R L I C G L T V TTAATTTATACATTAAACCCCTTTTGATGCTTTTAAATCGATTGATTGCGGGCTTACAGTT	10920
<b>Start of orf10</b>	
G V V Y G F A K I R *	<b>End of orf9</b>
M D L Q K L D K Y T C N G N L D A GGAGTTGTTTATGAGATTTCGAAAAATTAGATAAGTATACCTGTAATGGAATTTAGACGC	10980
P L V S I I I A T Y N S E L D I A K C L TCCACTTGTTTCAATAATCATTCGAACTTATAATTCGAACTTGATATAGCTTAAGTGT	11040
Q S V T N Q S Y K N I E I I I M D G G S GCAATCGGTAACATTAATCTTATAAGAATATTGAAATCATAATAATGGATGGAGGATC	11100
S D K T L D I A K S F K D D R I K I V S TTCTGATAAAACGCTTGATATGCAAAATCGTTTAAAGACGACCGAATAAAAATGATTTT	11160
E K D R G I Y D A W N K A V D L S I G D AGAGAAAGATCGTGAATTTATGATGCCTGGAATAAAGCAGTTGATTTATCCATTGGTGA	11220
W V A F I G S D D V Y Y H T D A I A S L TTGGGTAGCATTTATGGTTCAGATGATGTTTACTATCATACAGATGCAATTCGCTTCATT	11280
M K G V M V S N G A P V V Y G R T A H E GATGAAGGGGGTTATGGTATCTAATGGCGCCCTGTGGTTTATGGGAGGACAGCGCACGA	11340

14/96

G P D R N I S G F S G S E W Y N L T G F 11400  
 AGTCCCGGATAGAACATATCTGGATTTTCAGGCGAGTGAATGGTACAACTAACAGGATT  
 K F N Y Y K C N L P L P I M S A I Y S R 11460  
 TAAGTTTAATTATTACAAATGTAATTTACCATTGCCCATTTAGAGCGCAATATTCTCTG  
 D F F R N E R F D I K L K I V A D A D W 11520  
 TGATTTCTTCAGAAACGACGTTTGTATATTAATTAATAATTTGTTGCTGACGCTGATTG  
 F L R C F I K W S K E K S P Y F I N D T 11580  
 GTTTCGAGATGTTTCATCAAATGGAGTAAAGAGAGTACCTTATTTTATTATGACAC  
 T P I V R M G Y G G V S T D I S S Q V K 11640  
 GACCCCTATTGTTAGATAAGGATATGGTGGGGTTTCGACTGATATTTCTTCAAGTTAA  
 T T L E S F I V R K K N N I S C L N I Q 11700  
 AACTACGCTAGAAAGTTTCATTGTACGCAAAAGAATAATATATCCTGTTTAAACATACA  
 L I L R Y A K I L V M V A I K N I F G N 11760  
 GCTGATTCTTAGATATGCTAAAATTCGTGGTATGGTATGCGATCAAAAATATTTTGGCAA  
 N V Y K L M H N G Y H S L K K I K N K I 11820  
 TAATGTTTATAAAATTAATGCATAACGGGTATCATTCCTTAAAGAAAATCAAGAATAAAAT

## Start of orf11, End of orf10

M K I V Y I I T G L T C G G A E H L M T  
 \*  
 ATGAAGATTGTTTATATAATAACCGGGCTTACTTGTGGTGGAGCCGAACCTTATGACG 11880  
 Q L A D Q M F I R G H D V N I I C L T G 11940  
 CAGTTAGCAGACCAAAATGTTTATACGCGGGCATGATGTTAATATTATTGTTCACTAGTGT  
 I S E V K P T Q N I N I H Y V N M D K N 12000  
 ATATCTGAGGTAAAGCCAAACAAAATATTAATATTCATTATGTTAATATGGATAAAAAAT  
 F R S F F R A L F Q V K K I I V A L K P 12060  
 TTTAGAAGCTTTTTTAGAGCTTTATTTCAAGTAAAAAAAATAATTTGTCGCTTTAAAGCCA  
 D I I H S H M F H A N I F S R F I R M L 12120  
 GATATAATACATAGTCATATGTTTCATGCTAATATTTTATGTCGTTTATTAGGATGCTG  
 I P A V P L I C T A H N K N E G G N A R 12180  
 ATTCGAGCGGGTCCCTCATATGTACCGCACACAAAAATGAAGGTGGCAATGCAAGG  
 M F C Y R L S D F L A S I T T N V S K E 12240  
 ATGTTTGTATTCGACTGAGTGATTTTTTAGCTTCTATTACTACAAATGTAAGTAAAGAG  
 A V Q E F I A R K A T P K N K I V E I P 12300  
 GCTGTTCAAGAGTTTATAGCAAGAAAGGCTACACCTAAAAATAAAATAGTAGAGATTCCG  
 N F I N T N K F D F D I N V R K K T R D 12360  
 AATTTTATTAATACAAATAAATTTGATTTTGATATTAATGTGCAAGAAAGAAACGCGAGAT  
 A F N L K D S T A V L L A V G R L V E A 12420  
 GCTTTAATTTGAAAGACAGTACAGCAGTCTGCTCGCAGTAGGAAGACTTGTGTGAAGCA  
 K D Y P N L L N A I N H L I L S K T S N 12480  
 AAAGACTATCCGAACTTATTAATGCAATAATCATTTGATTCCTTCAAAAACATCAAAT  
 C N D F I L L I A G D G A L R N K L L D 12540  
 TGTAATGATTTATTTTGCTTATGTCTGGCATGGCGCATTAAGAAATAAATATTGGAT  
 L V C Q L N L V D K V F F L G Q R S D I 12600  
 TTGTTTGTCAATTAATCTTGTGGATAAAGTTTCTTCTTGGGGCAAGAGATGATAT

15/96

K E L M C A A D L F V L S S E W E G F G 12660  
 AAAGAATTAATGTCGTGCAGATCTTTTGTGTTGAGTTCGTGAGTGGGAAGGTTTGTG  
 L V V A E A M A C E R P V V A T D S G G 12720  
 CTCGTTGTGTCAGAAGCTATGGCGTGTGAACGTCGCCGTTGTGCTACCGATTCTGGTGA  
 V K E V V G P H N D V I P V S N H I L L 12780  
 GTTAAAGAAGTCGTGGACCTCATATGATGTTATCCCTGTCAGTAATCATATTCTGTTG  
 A E K I A E T L K I D D N A R K I I G M 12840  
 GCAGAGAAAATCGCTGAGACACTTAAAAATAGATGATAACGCAAGAAAAATATAGGTATG  
 K N R E Y I V S N F S I K T I V S E W E 12900  
 AAAAAATAGAAATATATTGTTTCCAATTTTCAATTAACAGCATAGTGAGTGAGTGGGAG

## End of orf11

R L Y F K Y S K R N N I I D \* 12960  
 CGCTTATATTTTAAATATTTCCAAGCGTAATAATATAATTGATGAAAAATATAAGTTTGT  
 CTCGTGAGTCAATAGTTTCTCTATGCTGTTTTTTTACTGGCTCCGATTTTCTTATATAG 13020  
 CTGGATTTTGTATATATACAGTATTAATCTGCTCAACTTCATCTAGACTACATTCACG 13080

## Start of gnd

M S K Q Q I 13140  
 CGCGCATCGCTCGCGCGGTGACTACACCTGACAGAGTATGTAATGCCAACACAGAT  
 G V V G M A V M G R N L A L N I E S R G 13200  
 CGGCGTCGTCGGTATGCGAGTGATGGGGCGCAACCTGGCGCTCAACATCGAAAGCCGCGG  
 Y T V S I F N R S R E K T E E V V A E N 13260  
 TTATACCGCTCCATCTTCAACCGCTCCCGCGAGAAACTGAGAAAGTTGTTGCCGAGAA  
 P D K K L V P Y Y T V K E F V E S L E T 13320  
 CCCGGATAAGAACTGGTTCCTTATTACACGGTGAAAGAGTTCTGTCAGTCTCTTGAAC  
 P R R I L L M V K A G A G T D A A I D S 13380  
 CCCACGTCGTATCTGTTAATGGTAAAGCAGGGGCGGGAAGTGTGCTGCTATCGATT  
 L K P Y L D K G D I I I D G G N T F F Q 13440  
 CCTGAACCGGTATCTGATAAAGCGACATCATTTATGATGGTGGCAACACCTTCTTCCA  
 D T I R R N R E L S A E G F N F I G T G 13500  
 GGACACTATCCGTCGTAACCGTGAACCTGCGCGGAAGGCTTTAACTTCATCGGTACCGG  
 V S G G E E G A L K G P S I M P G G Q K 13560  
 CGTGTCCGCGGTGAAGAGGGCGCCCTGAAAGGCCATCATATGCGAGGTGGCCAGAA  
 E A Y E L V A P I L T K I A A V A E D G 13620  
 AGAAGCGTATGAGCTGGTTGCGCCTATCCTGACCAAGATTGCTGCGGTTGCTGAAGATGG  
 E P C I T Y I G A D G A G H Y V K M V H 13680  
 CGAACCATGTATAACTGCGTCTGACGGTGGCGGCTCACTACGTGAAGATGGTGCA  
 N G I E Y G D M O L I A E A Y S L L K G 13740  
 CAACGGTATCGAATATGGCGATATGCAGCTGATTGCTGAAGCCTATCTCTGCTTAAAGG  
 G L N L S N E E L A T T F T E W N E G E 13800  
 CGGCTTAATCTGTCTAACGAAGAGCTGGCAACCACTTTTACCGAGTGAAGTGAAGGCGA  
 L S S Y L I D I T K D I F T K K D E E G 13860  
 GCTAAGTAGCTACTGATTGACATCAACAAAGACATCTTCAACAAAAAGATGAAGAGGG

16/96

K Y L V D V I L D E A A N K G T G K W T  
 TAAATACCTGGTTGATGTGATCCTGGACGAAGCTGCGAACAAGGCACCGGTAAATGGAC 13920  
 S Q S S L D L G E P L S L I T E S V F A  
 CAGCCAGAGCTCTCTGGATCTGGGTGAACCGCTGTCGCTGATCACCGAATCCGTATTTCGC 13980  
 R Y I S S L K D Q R I A A S K V L S G P  
 TCGCTACATCTCTTCTCTGAAAGACCAGCGCATTGCGGCATCTAAAGTGCTGTCTGGTCC 14040  
 Q A K L A G D K A E F V E K V R R A L Y  
 GCAGGCTAAACTGGCTGGTGATAAAGCAGAGTTCGTTGAGAAAGTCCGTCGCGCGCTGTA 14100  
 L G K I V S Y A Q G F S Q L R A A S D E  
 CCTGGGTAAATCGTCTCTTATGCCC AAGGCTTCTCTCAACTGCGTGCCGCGTCTGACGA 14160  
 Y N W D L N Y G E I A K I F R A G C I I  
 ATACAACTGGGATCTGAACTACGGCGAAATCGCGAAGATCTTCGCGCGGGCTGCATCAT 14220  
 R A Q F L Q K I T D A Y A E N K G I A N  
 TCGTGCGCAGTTCTCTGCAGAAAATTACTGACGCGTATGCTGAAAACAAAGGCATTGCTAA 14280  
 L L L A P Y F K N I A D E Y Q Q A L R D  
 CCTGTTGCTGGCTCCGTA CTCAA AATATCGCTGATGAATATCAGCAAGCGCTGCGTGA 14340  
 V V A Y A V Q N G I P V P T F S A A V A  
 TGTATGGCTTATGCTGTGCAGAACGGTATTCGGGTACCGACCTTCTCTGAGCGGTAGC 14400  
 Y Y D S Y R S A V L P A N L I Q A Q R D  
 CTACTACGACAGCTACCGTTCTGCGGTACTGCCGGCTAATCTGATT CAGGCACAGCGTGA 14460  
 Y F G A H T Y K R T D K E G V F H T G  
 TTACTTCGGTGCGCACACGTATAAACGCAC T GATAAAGAAGGTGTGTCCACACCG 14516

17/96

GTAACCAAGGGCGGTACGTGCATAAAATTTAATGCTTATCAAACTATTAGCATTAAAAA 60

Start of orf1

M N K E T V V S I I M P V V N  
TATATAAGAAATTCCTCAAAATGAAGAAAGAAACCGTTCAATAATTATGCCCGTTTACAAAT 120

G A K T I I S S V E S I I H Q S Y Q D F  
GGGGCCAAAACATATACTCATCAGTAGAATCAATTATACATCAATCTTATCAAGATTTT 180

V L Y I I D D C S T D D T F S L I N S R  
GTTTGTATATCATTGACGATTGTAGCACCAGTAGATACATTTTCATTAAATCAACAGTCGA 240

Y K N N Q K I R I L R N K T N L G V A E  
TACAAAAACAATCAGAAAAATAAGAATATTGCGTAACAGACAAATTTAGGTGTGCAGAA 300

S R N Y G I E M A T G K Y I S F C D A D  
AGTCGAAATATTGGAATAGAAATGGCCACGGGAAATATATTTCTTTTGTGATGCGGAT 360

D L W H E K K L E R Q I E V L N N E C V  
GATTTGTGGCAGAGAAAAATAGAGCGTCAAATCGAAGTGTAAATAATGAATGTGTA 420

D V V C S N Y Y V I D N N R N I V G E V  
GATGTGGTATGTTCTAATTATTATGTTATAGATAACAATAGAAATATTGTTGGCGAAGTT 480

N A P H V I N Y R K M L M K N Y I G N L  
AATGCTCCTCATGTGATAAATTATAGAAAAATGCTCATGAAAACTACATAGGGAAATTTG 540

T G I Y N A N K L G K F Y Q K K I G H E  
ACAGGAATCTATAATGCCAACAAATTGGGTAAGTTTATCAAAAAAGATTGGTCACGAG 600

D Y L M W L E I I N K T N G A I C I Q D  
GATTATTGTATGGGTGGAATAATTAAATAAAACAAATGGTGCTATTGTATTCAAGAT 660

N L A Y Y M R S N N S L S G N K I K A A  
AATCTGGCGTATTACATGCGTTCAAATAATTCACTACGGGTAATAAAATTAAGCTGCA 720

K W T W S I Y R E H L H L S F P K T L Y  
AAATGGACATGGAGTATATATAGAGAACATTACATTGTGCTCTTCCAAAAACATTATAT 780

Y F L L Y A S N G V M K K I T H S L L R  
TATTTTATTATATGCTTCAATGGAGTCATGAAAAATAACACATTCACATTAAAG 840

Start of orf2, End of orf1

R K E T K K \*  
V K S A A K L I F L F L F T  
AGAAAGGAGACTAAAAAGTGAAGTCAGCGGCTAAGTTGATTTTTTATTCCATTATTACAC 900

L Y S L Q L Y G V I I D D R I T N F D T  
TTATAGTCTCCAGTTGTATGGGGTTATCATAGATGATCGTATAACAAATTTTGATACAA 960

K V L T S I I I I F Q I F F V L L F Y L  
AGGTATTAACAGTATATAATATATTTTCAGATTTTTTTTGTGTTTATTATTATTCTAA 1020

T I I N E R K Q Q K K F I V N W E L K L  
CGATTATAAATGAAGAAAAACAGCAGAAAAATTTATCGTGAACTGGGAGCTAAAGTTAA 1080

I L V F L F V T I E I A A V V L F L K E  
TACTCGTTTCCCTTTGTGACTATAGAAATGCTGCTGTAGTTTATTCTTAAAGAAG 1140

G I P I F D D D P G G A K L R I A E G N  
GTATTCCTATATTGTATGATGATCCAGGGGGGCTAAACTTAGAATAGCTGAAGGTAATG 1200



18/96

G L Y I R Y I K Y F G N I V V F A L I I  
 GACTTTACATTAGATATATTAAGTATTTTGGTAATATAGTTGTGTTTGCATTAAATTTATTC 1260  
 L Y D E H K F K Q R T I I F V V F T T I  
 TTTATGATGAGCATAAATTCAAACAGAGGACCATCATATTTGTATATTTTACAACGATTG 1320  
 A L F G Y R S E L V L L I L Q Y I L I T  
 CTTTATTTGGTTATCGTTCTGAATTTGGTGTGCTCATTCTTCAATATATATTGATTACCA 1380  
 N I L S K D N R N P K I K R I I G Y F L  
 ATATCCTGTCAAAGGATAACCGTAATCTCTAAAATAAAAGAAATAATAGGGTATTTTAT 1440  
 L V G V V C S L F Y L S L G O D G E Q N  
 TGGTAGGGTTGTATGCTCGTTGTTTATCTAAGTTTAGGACAAGACGGAGAACAAAATG 1500  
 D S Y N N M L R I I N R L T I E Q V E G  
 ACTCATATAAATAATATGTTAAGGATAAATAATAGGTTAAACAAATAGACCAAGTTGAAGGTG 1560  
 V P Y V V S E S I K N D F F F P T P E L E  
 TTCCATATGTTGTTCTGAATCTATTAAGAACGATTCTTCTCCGACACAGAGTTAGAAA 1620  
 K E L K A I I N R I Q G I K H Q D L F Y  
 AGGAATTAAGCAATAATAAATAGAAATACAGGGAATAAAGCATCAAGACTTATTTATG 1680  
 G E R L H K Q V F G D M G A N F L S V T  
 GAGAACGGTTACATAACCAAGTATTTGGAGACATGGGAGCAAATTTTATCAGTTACTA 1740  
 T Y G A E L L V F F G F L C V F I I P L  
 CGTATGGAGCAGAACTGTAGTTTCTTTGGTTTCTCTGTGTATTCTATTCCTTTAG 1800  
 G I Y I P F Y L L K R M K K K T H S S I N  
 GGTATATATACCTTTTATCTTTTAAAGAGAATGAAAAAACCCATAGCTCGATAAAT 1860  
 C A F Y S Y I I M I L L Q Y L V A G N A  
 GCGCATTCATTATATATCATTTATGATTTTATGCAATACCTAGTAGGCTGGGAATGCAT 1920  
 S A F F F G P F L S V L I M C T P L I L  
 CGGCCCTCTTTTGGTCCTTTCTCTCGTATGATAATGTGTACTCTCTGATCTTAT 1980

Start of orf3

M K I S V I T V Y  
 L H D T L K R L S R N E N I S Y N C D L  
 TGCATGATACGTTAAAGAGATTATCAGAAATGAAAAATATCAGTTATAACTGTGACTTAT 2040

End of orf2

N N A E G L E K T L S S L S I L K I K P  
 AATAATGCTGAAGGGTTAGAAAAAACTTTAAGTAGTTTATCAATTTTAAAAATAAAACCT 2100  
 F E I I I V D G G S T D G T N R V I S R  
 TTTGAGATTATATAGTTGATGGCGGCTCTACAGATGGAACGAATCGTGTCTATAGAGA 2160  
 P T S M N I T H V Y E K D E G I Y D A M  
 TTTACTAGTATGAATATTACACATGTTTATGAAAAAGATGAAGGGATATATGATGCGATG 2220  
 N K G R M L A K G D L I H Y L N A G D S  
 AATAAGGGCCGAATGTTGGCCAAAGCGACTTAATACATTATTTAAACGCCGGCGATAGC 2280  
 V I G D I Y K N I K E P C L I K V G L F  
 GTAATTGGAGATATATATAAAAAATCAAAAGAGCCATGTTTGATTAAAGTTGGCCTTTTC 2340  
 E N D K L L G F S S I T H S N T G Y C H  
 GAAAATGATAAACTTCTGGGATTTTCTCTATAACCAATTCAAAATACAGGGTATTGTCAT 2400

19/96

Q G V I F P K N H S E Y D L R Y K I C A  
CAAGGGGTGATTTTCCCAAAGAATCATTCAGAATATGATCTAAGGTATAAAATATGTGCT 2460

D Y K L I Q E V F P E G L R S L S L I T  
GATTATAAGCTTATTCAAGAGGTGTTTCTCGAAGGGTTAAGATCTCTATTCTTTGATTACT 2520

S G Y V K Y D M G G V S S K K R I L R D  
TCGGGTATGTAAAAATATGATATGGGGGAGTATCTTCAAAAAAAGAATTTTAAAGAGAT 2580

K E L A K I M F E K N K K N L I K F I P  
AAAGAGCTTGCCAAAATATGTGTTGAAAAATAAAAAAACCTTATTAAGTTTATTTCCA 2640

I S I I K I L F P E R L R R V L R K M Q  
ATTTCAATAATCAAAATTTTATTCCTGAACGTTTAAAGAAGATATGCGGAAAATGCAA 2700

Start of orf4 End of orf3

Y I C L T L F F M K N S S P Y D N E \*  
M I M N K I  
TATATTTGTCTAACTTTATTCTTCATGAAGAATAGTTCACCATATGATAATGAATAAAAT 2760

K K I L K F C T L K K Y D T S S A L G R  
CAAAAAATACTTAAATTTTGCACCTTTAAAAAATATGATACATCAAGTGCCTTTAGGTAG 2820

E Q E R Y R I I S L S V I S S L I S K I  
AGAACAGGAAAGGTACAGGATTATATCTTGTCTGTTATTCAAGTTTGAATAGTAAAAAT 2880

L S L L S L I L T V S L T L P Y L G Q E  
ACTCTACACTTTCTCTTATTAAGTAAAGTTTAACTTTACCTTATTTAGGACAAGA 2940

R F G V W M T I T S L G A A L T F L D L  
GAGATTGGTGATGGGATGACTATTACCAGTCTTGGTGCTGCTGACATTTTGGGACTT 3000

G I G N A L T N R I A H S F A C G K N L  
AGGTATAGGAAATGCATTAACAACAGGATCGCACATTCATTTGCGTGTGGCAAAAATTT 3060

K M S R Q I S G G L T L L A G L S F V I  
AAAGATGAGTCGGCAAAATAGTGGTGGGCTCACTTTTGTGCTGGGATTATCGTTTGTCTAT 3120

T A I C Y I T S G M I D W Q L V I K G I  
AACTGCAATATGCTATATTACTTCTGGCATGATTGATTGGCAACTAGTAATAAAAGGTAT 3180

N E N V Y A E L Q H S I K V F V I I F G  
AAACGGAATGTGTATGCAGAGTTACAACACTCAATTAAGTCTTTGTAATCATATTTGG 3240

L G I Y S N G V Q K V Y M G I Q K A Y I  
ACTTGGAAATTTATTCAAATGTGTGCAAAAAGTTTATATGGGAATACAAAAGCCTATAT 3300

S N I V N A I F I L L S I I T L V I S S  
AAGTAATATTGTTAATGCCATATTTATATTGTTATCTATTACTCTAGTAATATCGTC 3360

K L H A G L P V L I V S T L G I Q Y I S  
GAAACTACATGCGGGACTACCAGTTTAAATGTGACAGCTCTGGTATTCAATACATATC 3420

G I Y L T I N L I I K R L I K F T K V N  
GGGAATCTATTAAACAATTAATCTTATTATAAAGCGATTATAAAGTTTACAAAAGTTAA 3480

I H A K R E A P Y L I L N G F F F F I L  
CATACATGCTAAAAGAGAAGCTCCATATTTGATATTAAACGGTTTTCCTTTTATTTT 3540

Q L G T L A T W S G D N F I S I T L G  
ACAGTTAGGCACTTGGCAACATGGAGTGGTGATAAATTTATAATATCTATAACATTGGG 3600

20/96

V T Y V A V F S I T Q R L F Q I S T V P 3660  
TGTACTACTATGTTGCTGTTTTCAGCATTACACAGAGATTATTTCAAATATCTACGGTCCC  
L T I Y N I P L W A A Y A D A H A R N D 3720  
TCTTACGATTATTAACATCCCCTTATGGGCTGCTTATGCAGATGCTCATGCACGCAATGA  
T Q F I K K T L R T S L K I V G I S S F 3780  
TACTCAATTATTAATAAAAGACGCTCAGAACATCATTGAAAATAGTGGGTATTTTCATCATT  
L L A F I L V V F G S E V V N I W T E G 3840  
CTTATTGGCCTTCATATTAGTAGTGTTCGGTAGTGAAGTCGTTAATATTGGACAGAAGG  
K I Q V P R T F I I A Y A L W S V I D A 3900  
AAAGATTACAGGTACCTCGAACATTATAATAGCTTAGCTTTATGGTCTGTTATTGATGC  
F S N T F A S F L N G L N I V K Q Q M L 3960  
TTTTTCGAATACATTGCAAGCTTTTAAATGGTTTGAACATAGTTAAACAACAAATGCT  
A V V T L I L I A I P A K Y I I V S H F 4020  
TGCTGTTGTAAACATTGATATTGATCGCAATTCAGCAAAATACATCATAGTTAGCCATTG  
G L T V M L Y C F I F I Y I V N Y F I W 4080  
TGGGTTAAGCTTATGTTGCTGCTGCTTCAATTTTATATATATATGTAATTAATTAATG

Start of orf5, End of orf4  
M K M

Y K C S F K K H I D R Q L N I R G \* 4140  
GTATAAATGTAAGTTTAAAAAACATATCGATAGCAGTTAAAAATAAGAGGATGAAAAATG  
K Y I P V Y Q P S L T G K E K E Y V N E 4200  
AAATATATACGAGTTTACCAACCGTCAATTGACAGGAAGAAAGAAATATGTAAATGAA  
C L D S T W I S S K G N Y I Q K F E N K 4260  
TGTCTGGAAGCTCAAGCTGGATTTCATCAAAAGGAAGTATATTCAGAAAGTTTGAAATAAA  
F A E Q N H V Q Y A T T T V S N G T V A L 4320  
TTTGGGGAAACAAAACCATGTGCAATATGCAAGTACGTAAAGTAATGGAAGCGTTGCTGCTT  
H L A L L A L G I S E G D E V I V P T L 4380  
CATTTAGCTTTTGTAGCGTTAGGTATATCGGAAGGAGATGAAGTTATGTTTCCACACTG  
T Y I A S V N A I K Y T G A T P I F V D 4440  
ACATATATAGCATCAGTTAATGCTATATAAATACACAGCAGCAGCCCGCATTTTCGTTGAT  
S D N E T W Q M S V S D I E Q K I T N K 4500  
TCAGATATATGAAGCTTGGCAATGTCTGCTTAGTGACATAGAACAAAAATCACTATATATA  
T K A I M C V H L Y G H P C D M E Q I V 4560  
ACTAAAGCTATATATGCTGTGCTCAATTTATACGGAGATGCAATGTGATATGGAACAAATGTA  
E L A K S R N L F V I E D C A E A F G S 4620  
GAAGCTGGCAAAAGCTAGAAATTTGTTTGTAAATGGAAGTTGGCGTGAAGCGTTTGGTCTT  
K Y K G K Y V G T F G D I S T F S F F G 4680  
AAATATAAGGTAATATGTGGGAACATTTGGAGATATTTTCTACTTTTATGCTTTTGGAA  
N K T I T T G E G G M V V T N D K T L Y 4740  
AATAAACATTAATACAGGTGAAGGTGGAAATGCTGTGCAAGATGACAAAACACTTTAT  
D R C L H F K G Q G L A V H R Q Y W H D 4800  
GACCGTGTGTTACATTTTAAAGGCAAGGATTAGCTGTGATACAGGCAATATTTGCAATGAC  
V I G Y N Y R M T N I C A A I G L A Q L 4860  
GTTATAGGCTAGCAATATATAGGATGACAAATATCTGCGCTGCTATAGGATTAGCCGATGTA

21/96

E Q A D D F I S R K R E I A D I Y K K N	4920
GAACAAGCTGATGATTTTATATACGAAAAAGCTGAATTTGCTGATATTTATAAAAAAAT	
I N S L V Q V H K E S K D V F H T Y W M	4980
ATCAAGAGTCTGTGACAAAGTCCACAAAGGAAGTAAAGATGTTTTTCACACTTATTTGGATG	
V S I L T R T A E E E R E E L R N H L A D	5040
GTCTCAATTTCAACTAGGACCCGACGAGGAAGAGAGGAATTAAGCAATCACTTGCAGAT	
K L I E T R P V F Y P V H T M P M Y S E	5100
AACTCATCGAAGCAAGGCCAGTTTTTACCCTGTCCACAGATGCCAATCTAGCTCGGAA	
K Y Q K H P I A E D L G W R G I N L P S	5160
AAATATCAAAAGCAACCTATAGCTGAGGATCTTGGCTTGGCGTGAATTAATTTAGCTAGT	
F P S L S N E Q V I Y I C E S I N E F Y	5220
TTCCCCAGCCTATCGAATGAGCAAGTTATTTATATTTTGAATCTATTAACGAATTTTAT	
<div style="display: flex; justify-content: space-between; margin-top: 20px;"> <div> <p><b>End of orf5</b></p> <p>S D K *</p> </div> <div> <p><b>Start of orf6</b></p> <p>M K I A L N S D</p> </div> </div>	
AGTGTATAAATCTGCTAAATATTTCTAAAGGTCATCTCATGAAAAATTCGATTCAGAT	5280
G F Y E W G G G I D F I K Y I L S I L E	5340
GGATTTTACGAGTGGGGCGGTGGGAATTTGATTTTATTAATATATATCTGTCAATATTAGAA	
T K P E I C I D I L L P R N D I H S L I	5400
ACGAAAACCGAAATATGTATCGATATCTTTTACCGAGAAATGATATACATTTCTCTTATA	
R E K A F P P F K S I L K A I L K R E R P	5460
AGAGAAAAAGCATTTCTCTTTTAAAGATATATTAAGCAATTTTAAAGAGGGAAAGGCCT	
R W I S L N R F N E Q Y Y R D A F T Q N	5520
CGATGGATTTTCATTAAATAGATTTAATGAGCAATACTATAGAGATGCCTTTACACAAAAAT	
N I E T N L T F I K S K S S A F Y S Y F	5580
AATATAGAGACGAATCTTACCTTTATTAAGAGTAAGAGCTCTGCCTTTTATTCATTTTT	
D S S D C D V I L P C M R V P S G N L N	5640
GATAGTAGCGGATTTGATGTTATTTCTCTCTTGCGCTGCTCTTCGGGAAATTTGAAT	
K K A W I G Y I Y D F Q H C Y Y P S F F	5700
AAAAAAGCATGGATGGTTATATTTTGACTTCAACACTGTACTATCTTCATTTTTT	
S K R E I D Q R N V F F K L M L N C A N	5760
AGTAAGCGAGAAATAGATCAAAGGAATGTGTTTTTAAATGTAGTCTCAATTCGCGTAAC	
N I I V N A H S V I T D A N K Y V G N Y	5820
AAATATTATTGTTAATGCACATTCAGTTATTACCGATGCAAAATAAATATGTTGGGAATTAT	
S A K L H S L P F F S P C P Q L K W F A D	5880
CTCGAAAAACTACATTTCTTCCATTTAGTCCATGCCCTCAATTAATGGTTTCGCTGAT	
Y S G N I A K Y N I D K D Y F I I C N Q	5940
TACTCTGGTAATTTGGCAAATATAATATTGACAAGGATTTATTTATAATTTGCAATCAA	
F W K H K D H A T A F R A F K I Y T E Y	6000
TTTGGAAACATAAGATCATGCAACTGCTTTTAGGGCATTTAAAAATTTATACTGAATAT	
N P D V Y L V C T G A T Q Q D Y R F P G Y	6060
AATCTGATGTTTTATTTAGTATGACGGGAGCTACTCAAGATTATCGATTCCCTGGATAT	
F N E L M V L A K K L G I E S K I K I L	6120
TTTAAATGAATGATGGTTTTGGCAAAAAGCTCGGAATTTGAATCGAAAAATTAAGATATTA	

22/96

G H I P K L E Q I E L I K N C I A V I O  
GGGCATATACCTAAACTTGAACAAATTGAATTAATCAAAAAATGCATTGCTGTAATACAA 6180

P T L F E G G P G G G V T F D A I A L G  
CCAACTTATTTGAAGCGGGCTTGGAGGGGGGTAACTTTGACGCTATTGCGATTAGGG 6240

K K V I L S D I D V N K E V N C G D V Y  
AAAAAGTTATACCTATCTGACATAGATGTCATAAAGAAAGTTAATTGCGGTGATGATATAT 6300

F F Q A K N H Y S L N D A M V K A D E S  
TTCTTTTCAGGCAAAAAACCATTATTCATTAATGACGCGATGGTAAAAGCTGATGAATCT 6360

K I F Y E P T T L I E L G L K R R N A C  
AAAATTTTTATGAACCTACAACCTCTGATAGAATTGGGTCTCAAAAGACGCAATGCGTGT 6420

## End of orf6

A D F L L D V V K Q E I E S R S \*  
GCAGATTTCTCTTTAGATGTTGTGAAACAAGAAATTGAATCCCGATCT TAATATATTCAA 6480

## Start of orf7

M T K V A L I T G V T G Q D G S Y  
GAGGTATATAATGACTAAAGTCGCTCTTTATACAGGTGTAACCTGGACAAGATGGAATCTTA 6540

L A E F L L D K G Y E V H G I K R R A S  
TCTAGCTGAGTTTTTCTGTGATAAAGGGTATGAAGTTTCATGGTATCAAAACGCCGAGCCTC 6600

S F N T E R I D H I Y Q D P H G S N P N  
ATCTTTTAATACAGAACGATAGACCATATTTATCAAGATCCACATGGTCTTAACCCAAA 6660

F H L H Y G D L T D S S N L T R I L K E  
TTTTCACTTGCACTATGGAGATCTGACTGATTCATCTAACCTCACTAGAATTCATAAGGA 6720

V Q P D E V Y N L A A M S H V A V S F E  
GGTACAGCCAGATGAAGTATATAATTTAGCTGCTATGAGTCACGTAGCAGTTTCTTTTGA 6780

S P E Y T A D V D A I G T L R L L E A I  
GTCTCCAGAATATACAGCCGATGTCGATGCAATTGGTACATTACGTTTACTGGAAGCAAT 6840

R F L G L E N K T R F Y Q A S T S E L Y  
TCGCTTTTTAGGATTGGAAAAACAAACGCGTTTCTATCAAGCTTCAACCTCAGAAATTATA 6900

G L V Q E I P Q K E S T P F Y P R S P Y  
TGGACTTGTTTCAGGAAATCCCTCAAAAGAAATCCACCCCTTTTTATCTCTCGTCCCTCTTA 6960

A V A K L Y A Y W I T V N Y R E S Y G I  
TGCAGTTGCAAAACTTTACGCATATTGGATCACGGTAAATTCAGAGTCATATGGTAT 7020

Y A C N G I L F N H E S P R R G E T F V  
TTATGTCATGTAATGGTATATTGTTCAATCATGAATCTCCACGCCGCTGGAGAAACGTTTGT 7080

T R K I T R G L A N I A Q G L E S C L Y  
AACAAGGAAAAATTACTCGAGGACTTGCAAATATTGCACAGGCTTGAATCATGTTTGTGA 7140

L G N M D S L R D W G H A K D Y V R M Q  
TTTAGGGAATATGGATTCTGTTACGAGATTGGGGACATGCAAAAGATTATGTTAGAAATGA 7200

W L M L Q Q E Q P E D F V I A T G V Q Y  
ATGGTTGATGTACAAACAGGAGCAACCCGAAGATTTTGTGATTGCAACAGGAGTCCCAATA 7260

S V R Q F V E M A A A Q L G I K M S F V  
CTCAGTCCGTCAGTTTGTGCAAAATGGCAGCAGCACAACTTGGTATTAGATGAGCTTTGT 7320

23/96

G K G I E E K G I V D S V E G Q D A P G 7380  
 TGGTAAAGGAATCGAAGAAAAGGCATTGTAGATTTCGGTTGAAGGACAGGATGCTCCAGG  
 V K P G D V I V A V D P R Y F R P A E V 7440  
 TGTGAAACCAGGTGATGTCATTGTTGCTGTTGATCTCGTATTTCGACCCAGCTGAAAT  
 D T L L G D P S K A N L K L G W R P E I 7500  
 TGATACTTTCCTGGAGATCCGAGCAAAGCTAATCTCAAACCTGGTGGAGACCAGAAAT  
 T L A E M I S E M V A K D L E A A K K H 7560  
 TACTCTTGCTGAAATGATTTCTGAAATGGTTGCCAAAGATCTTGAAGCCGCTAAAAACA

Start of orf8, End of orf7

M M M N K

S L L K S H G F S V S L A L E \* 7620  
 TTCTCTTTTAAATCGCATGGTTTTCTCTGAAGCTTAGCTCTGGAATGATGTAATAAG  
 Q R I F I A G H Q G M V G S A I T R R L L 7680  
 CAACGTATTTTATGCTGGTCACCAAGGAATGGTTGGATCAGCTATTACCCGACGCCCTC  
 K Q R D D V E L V L R T R D E L N L L D 7740  
 AAACAACGTGATGATGTTGAGTTGGTTTTACGTACTCGGGATGAATTGAACCTGTTGGAT  
 S S A V L D P F S S Q K I D Q V Y L A A 7800  
 AGTAGCGCTGTTTTGGATTTTTTTCTTCACGAAATCGACCGAGTTTATTTGGCAGCA  
 A K V G G I L A N S S Y P A D F I Y E N 7860  
 GCAAAAGTCGGAGGTATTTAGCTAACAGTTCTTATCCTGCCGATTTTATATATAGAAT  
 I M I E A N V I H A A H K N N V N K L L 7920  
 ATAATGATAGAGCGCAATGTCATTTCATGCTGCCACAAAAAATGTAATAAATGCTT  
 F L G S S C I Y P K L A H Q P I M E D E 7980  
 TTCTTCGGTTTCGTCGTGATTTATCTTAAGTTAGCACACCAACCGATTATGGAAGACGAA  
 L L Q G K K L E P T N E P Y A I A K I A G 8040  
 TTATTACAAAGGAACTTGAGCCAACAAATGAACCTTATGCTATCGCAAAAATTCGAGGT  
 I K L C E S Y N R Q F G R D Y R S V M P 8100  
 ATTAATATGATGAATCTTATAACCGTCAGTTTGGCGTGATTACCGTTCAGTAATGCCA  
 T N L Y G P N D N F H P S N S H V I P A 8160  
 ACCAATCTTTATGGTCCAAATGACAATTTTCATCCAAGTAATTCATGATGATTCGGCG  
 L L R R F H D A V E N N S P N V V W G 8220  
 CTTTTCGCCCGCTTCATGATGCTGTGAAAAACAATTCCTCGAATGTTGTTGTTGGGGA  
 S G T P K K R E F L H V D D M A S A S I Y 8280  
 AGTGGTCTCCAAAGCGGAATCTTACATGTAGATGATATGGCTTCGCAAGCATTTAT  
 V M E M P Y D I W Q K N T K V M L S H I 8340  
 GTCATGGAGATGCCATACGATATATGGCAAAAAATACTAAAGTAATGTTGCTTCATATC  
 N I G T G I D C T I C E L A E T I A K V 8400  
 AATATGGAACAGGTATTGACTGCACGATTTGTGAGCTTCGGAAACAATAGCAAAAGTT  
 V G Y K G H I T F D T T K P D G A P R K 8460  
 GTAGGTTATAAAGGCCATATTACGTTTCGATACACAACAGCCGATGGAGCCCCCTCGAAAA  
 L L D V T L L H Q L G W N H K I T L H K 8520  
 CTACTTGATGTAACGCTTCTTCATCAACTAGGTTGGAATCATAAAATTACCTCTTCAAG

24/96

		End of orf8
G L E N T Y N W F L E N Q L Q Y R G *		
GGTCTTGAATAACATACATACTGGTTCTTGAAACCAACTTCAATATCGGGG TAATAA	8580	
<b>Start of orf9</b>		
M F L H S Q D F A T I V R S T P L I S I		
TGTTTTACATTCCCAAGACTTTGCCACAATTGTAAGGTCTACTCCTCTATTTCATATAG	8640	
D L I V E N E F G E I L L G K R I N R P		
ATTTGATTGGAAAACGAGTTTGCGGAAATTTTGCTAGAAAACGAATCAACCGCCGCG	8700	
A Q G Y W F V P G G R V L K D E K L Q T		
CACAGGGCTATTGGTTCGTCTCGTGGTAGGGTGTTGAAAGATGAAAAATGCAGACAG	8760	
A F F E R L T E I E L G I R L P L S V G K		
CTTTGAACGATTGACAGAAATTGAACTAGGAATTCGTTGCCTCTCTGTGGGTAAGT	8820	
F Y G I W Q H F Y E D N S M G G D F S T		
TTTATGGTATCTGGCGACACTTCTACGAAACAAATAGTATGGGGGGAGACTTTTCAACGC	8880	
H Y I V I A F L L K L Q P N I L K L P K		
ATTATATAGTTATAGCATTCCTCTCTAAATTACAAACCAACATTTTGAAATTACCGAAGT	8940	
S Q H N A Y C W L S R A K L I N D D V		
CACAACATAATGCTTATGTCTGGCTATCGCGAGCAAAGCTGATAAATGATGACGATGTGC	9000	
H Y N C R A Y F N N K T N D A I G L N		
ATTATAATTGTGCGCATATTTTAAACAATAAACAATGATGCGATTGGCTTAGATAATA	9060	
<b>Start of orf10      End of orf9</b>		
M S D A P I I A V V M A G G T G S		
K D I I C L M R Q *		
AGGATATAATATGCTCTGATGCGCAAATAATTGCTGTAGTTATGGCCGGTGGTACAGGCAG	9120	
R L W P L S R A E Y C Q F L Q L S G D		
TCGCTCTTTGGCCACTTTCTCGTGAECTATATCCAAAGKAGCTTTTACACTCTSGGTGA	9180	
N T L L Q T T L L R L S G L S C Q K P L		
TAAACACCTTGTACAAACGACTTTGCTACGACTTTCAGGCCTATCATGTGCAAAACCATT	9240	
V I T N E Q H R F V V A E Q L R E I N K		
AGTGATAACAAAGACAGCATCGCTTTGTGTGGCTGAAACAGTTAAGGGAATAATAATA	9300	
L N G N I I L E P C G R N T A P A I A I		
ATTAAATGGTAATATTATCTAGAAACCATCGCGGCGAAATACTGCACCAGCAATAGCGAT	9360	
S A F H A L K R N P Q E D P L L L V L A		
ATCTCGGTTTACGCGTTAAACCGTAATCCTCAGGAAGATCCATTGCTTCTAGTCTTGC	9420	
A D H V I A K E S V F C D A I K N A T P		
GGCAGACACCGTTATAGCTAAAGAAGTGTTFTTCTGTATGCTATAAAAATGCAACTCC	9480	
I A N Q K I V T F G I I P E Y A E T G		
CATCGCTAATCAAGTAAAAATTGAACGTTTGGAAATTATACCAGAATATGCTGAAACTGG	9540	
Y G Y I E R G E L S V P L Q G H E N T G		
TTATGGGTATTTAGAGAGGTGAACATCTGTACCCTTCAAGGCATGAAAACTACTGG	9600	
F Y Y V N K F V E K P N R E T A E L Y M		
TTTATTATATGATAAATAAGTTTGTCGAAAAGCCTAATCGTGAACCGCAGAAATTGTATAT	9660	
T S G N H Y W N S G A I F M F K A S V Y L		
GACTTCGTGTAATCACTATTGGAATAGTGGAAATTCATGTTTAAAGCATCTGTTTATCT	9720	

25/96

E E L R K F R P D I Y N V C E Q V A S S  
 TGAGGAATTGAGAAAAATTAGACCTGACATTTACAAATGTTGTGAACAGGTGCGCTCATC 9780  
 S Y I D L D F I R L S K E O F O D C P A  
 CTCATACATTGATCTAGATTTTATTCGATTATCAAAAAGAACAATTTCAAGATTGTCCCTGC 9840  
 E S I D F A V M E K T E K C V V C P V D  
 TGAATCTATTGATTTTGTCTGAATGGAAAAACAGAAAAATGTGTTGTATGCCCTGTGTA 9900  
 I G W S D V G S W Q S L W D I S L K S K  
 TATTGGTTGGAGTGACGTTGGATCTTGGCAATCGTTATGGACATTAGTCTAAAAATCGAA 9960  
 T G D V C K G D I L T Y D T K N N Y I Y  
 AACAGGAGTGATGTAAAGGTGATATATTAACCTATGATACTAAGAATAATTATATCTA 10020  
 S E S A L V A A I G I E D M V I V Q T K  
 CTCGTAGTCAGCGTTGGTAGCCGCCATTGGAATGGAAGATATGGTTATCGTGCAAACTAA 10080  
 D A V L V S K K S D V Q H V K K I V E M  
 AGATGCCGTTCTTGTGTCTAAAAAGAGTGATGTACAGCATGTAAAAAAAATAGTCGAAAT 10140  
 L K L Q Q R T E Y I S H R E V F R P W G  
 GCTTAAATTGCGACCAACGTACAGAGTATATTAGTCATCGTGAAGTTTTCGGACCATGGGG 10200  
 K F D S I D Q G E R Y K V K K I I V K P  
 AAAAATTGTGATTCGATTGACCAAGGTGAGCGATACAAAGTCAGAAAAATATTGTGAAACC 10260  
 G E G L S L R M H H H R S E H W I V L S  
 TGGTGAGGGGCTTCTTTAAGGATGCATCACCATCGTTCTGAACATTGGATCGTGCCTTTC 10320  
 G T A K V T L G D K T K L V T A N E S I  
 TGGTACAGCAAAGTAACCTTGGCGATAAAACTAAACTAGTCACCGCAAAATGAATCGAT 10380  
 Y I P L G A A Y S L E N P G I I P L N L  
 ATACATTCCCTTGGCGACGCTATAGTCTTGAGAATCCGGGCATAATCCCTCTTAATCT 10440  
 I E V S S G D Y L G E D D I I R Q K E R  
 TATTGAAGTCAGTTCAAGGGATTATTTTGGGAGAGGATGATATTATAAGACAGAAAGAACG 10500  
 End of orf10 Start of orf11  
 Y K H E D \* M K S L T C F K A Y D I R  
 TTACAAACATGAAGATTAAACATATGAATCTTTAACCTGCTTTAAAGCCATGATATTTCG 10560  
 G K L G E E L N E D I A W R I G R A Y G  
 CGGGAATTTAGCGGAAGAACTGAATGAAGATATTGCTTGGCGCATTTGGGCGTGCTTATGG 10620  
 E F L K P K T I V L G G D V R L T S E A  
 CGAATTTCTCAAAACCGAAAAACCATGTTTTAGCGGTTGATGTCGCGCTCACCGCGAAGC 10680  
 L K L A L A K G L Q D A G V D V L D I G  
 GTTAAACTGGGCGTTGCGAAAGGTTTACAGGATGCGGGCGTCGATGTGCTGGATATCGG 10740  
 M S G T E E I Y F A T F H L G V D G G I  
 TATGTCGCGACCGAAGAGATCTATTTGCGCACGTTCCATCTCGGAAGTGGATGGCGGAC 10800  
 E V T A S H N P M D Y N G M K L V R E G  
 CGAAGTTACCGCAGCCATAACCCGATGGATTACAACCGCATGAAGCTGGTGGCGGAAGG 10860  
 A R P I S G D T G L R D V Q R L A E A N  
 GGCTCGCCCGATCAGCGGTGATACCGGACTGCGCGATGTCCAGCGTCTGGCAGAAGCCAA 10920  
 D F P P V D E T K R G R Y Q Q I N L R D  
 TGACTTCCCTCTGTCGATGAAACCAAACGTGGTTCGCTATCAGCAAAATCAATCTGCGTGA 10980



26/96

A Y V D H L F G Y I N V K N L T P L K L  
 CGCTTACGTTGATCACCTGTTTCGGTTATATCAACGTCAAAAACCTCACGCCGCTCAAGCT 11040  
 V I N S G N G A A G P V V D A I E A R F  
 GGTGATCAACTCCGGGAACGGCGCAGCGGGTCCGGTGGTGGACGCCATTGAAGCCCGATT 11100  
 K A L G G A P V E L I K V H N T P D G N F  
 TAAGACCCCTCGGCGACCGGTGAATTAAATCAAAGTACACAACACGCCGGACGGCAATT 11160  
 P N G I P N P L L P E C R D D T R N A V  
 CCCCAACGGTATTCCTAACCCGCTGCTGCCGGAATGCCGCGACACACCCGTAATGCGGT 11220  
 I K H G A D M G I A F D G D F D R C F L  
 CATCAAAACACGGCGCAGGATATGGGCATTGCGCTTTGATGGCGATTGAGCCGCTGTTTCCT 11280  
 F D E K G Q F I E G Y Y I V G L L A E A  
 GTTTGACGAAAAAGGCGAGTTTATCGAGGCTACTACATTGTGCGCCTGCTGGCAGAAG 11340  
 F L E K N P G A K I I H D P R L S W N T  
 GTTCCCTCGAAAAATCCGCGCGAAGATCATCCACGATCCAGCTCTCTCCTGGAAACAC 11400  
 V D V V T A A G G T P V M S K T G H A F  
 CGTTGATGTGTGACTGCGCGAGCGCGCACCCCGGTAATGTGCAAAACCGGACACGCTT 11460  
 I K E R M R K E D A I Y G G E M S A H H  
 TATTAAGAAGCGTATGCGCAAGGAAGACCCCATCTACGGTGGCGAAATGAGCGCTACCA 11520  
 Y F R D F A Y C D S G M I P W L L V A E  
 TTACTTCGGTATTTTCGCTTACTGCGACAGCGGCATGATCCCGTGGCTGCTGCTGCGCA 11580  
 L V C L K G K T L G E M V R D R M A A F  
 ACTGGTGTGCTGAAAGAAAAACGCTGGGCGAAATGGTGGCGACACCGGATGCGCGCGTT 11640  
 P A S G E I N S K L A Q P V E A I N R V  
 TCCGGCAAGCGGTGAGATCAACAGCAAACCTGGCGCAACCCGTTGAGGCAATTAATCGCGT 11700  
 E Q H F S R E A L A V D R T D G I S M T  
 GGAACAGCATTTTAGCCGCGAGGCGCTGGCGGTGGATCGCACCGATGGCATCAGCATGAC 11760  
 F A D W R F N L R S S N T E P V V R L N  
 CTTTGGCGACTGGCGCTTTAACTGCGCTCCTCCAACACCGAACCAGTGGTGGCGGTGAA 11820  
 V E S R G D V K L M E K K K T K A L L K L  
 TGTGGAATACCGCGGTGATGTAAGCTAATGGAAAAGAAAACCTAAAGCTCTTCTTAATTT 11880

## End of orf11

L S  
 GCTAAGTGAGTGATATTATTACATTAATCATTAAGCGTATTTAAGATTATATTAAGTAAT 11940  
 GTTATTCGGTATATGATGAATATGTGGCTTTTTTATGTATAACGACTATACCGCAACT 12000

## Start of H-repeat

TTATCTAGGAAAAGATTAAAGAAAATAAAGTTTGTACTGACCAATTTGCATTTCACGTC 12060  
 ACGATIGAGACGTTCCCTTTGCTTAAGACATTTTTTCATCGCTTATGTAATAACAATGTG 12120  
 CCTTATATAAAAAGGAGAAACAAATGGAACCTTAAATAAATTGAGACAATAGATTTTTATT 12180  
 ATCCCTGTTTACGATATTATAGCCAAAGTTGTATCCTGCATCAGTCCTGCAATATTCAC 12240  
 GAGTGCTTTGTTAACTGAATACATGTCTGCCATTTTCAGATGATAACGACGTCATCGCA 12300  
 ATTGATGGTAAACACTTCGCGACACTTATGACAAGAGTCGTGCGAGAGGATGGTTTCAT 12360

27/96

GTCATTAGTGCCTTCAGCAATGCACAGTCTGGTCTCGGATAGATCAAGACGGATGAGA 12420

AACCTTAATGCGTTACAGCTTATTCATGAACCTTCTAAAATGATGGGTATTAAAGGAAAAA 12480

TAATCATAACTGATGCGATGGCTTGCCAGAAAAGATATTGACAGAGAAGATATAAAACAGA 12540

GATGTGATTATTATTCGCTGTAAAAGGAAATAAGAGTCGGCTTAATAGAGTCTTTGAGG 12600

AGATATTTACGCTGAAAAGAAATAAATAATCCAAAACATGACAGTTACGCAATTAGTGAAA 12660

AGAGGCACGGCAGAGACGATGTCGGTCTTCATATTGTTTGGATGCTCCTGATGAGCTTA 12720

TTGATTTACGTTTGAATGGAAAAGGGCTGCAGAATTTATGAATGGCAGTCCACTTTCTCT 12780

CAATAATAGCAGAGCAAAAGAAAGATCCGAAATGACGATCAAAATATTATATTAGATCTG 12840

CTGCTTTAACCCAGAGAAAGTTCGCCACAGTAAATCGAAATCACTGGCGCATGGAGAATA 12900

AGTTCACAGTAGCCTGATGTGGTAATGAATGAAATCGACTATAATATAAGAAGCGGAGT 12960

TGCATTCGAATGATTTTCTAGAATGCGGCACATCGCTATTAATATCTGACAATGATAATG 13020

TATTTCAAGGCAGGATTATCATGTAAGATGCGAAAAGCAGTCAAGACAGAAACTTCCTAG 13080

**End of the H-repeat**

CGTCAGGCATTGCAGCGTGGCGGCTTTCATAATCTTGCAT TGGTTTGTATAAGATATTTC 13140

**Start of orf12**

M N L Y G I F G A G S Y G R E

TTTGGAGATGGGAAAATGAATTTGTATGGTATTTTGTGCTGGAAGTTATGGTAGAGAA 13200

T I P I L N Q Q I K Q E C G S D Y A L V

ACAATACCCATTCTAAATCAACAAATAAAGCAAGAAATGGGTTCTGACTATGCTCTGGTT 13260

F V D D V L A G K K V N G F E V L S T N

TTTGTGGATGATGTTTGGCAGGAAAGAAAGTTAATGGTTTGAAGTGCTTTCAACCAAC 13320

C F L K A P Y L K K Y F N V A I A N D K

TGCTTTCTAAAGCCCCATTATTAAAAAGTATTTTAATGTGCTATGCTAATGATGAAG 13380

I R Q R V S E S I L L H G V E P I T I K

ATACGACAGAGAGTGCTGAGTCAATATTATTACACGGGCTTGAACCAATAACTATAAAA 13440

H P N S V V Y D H T M I G S G A I I S P

CATCCAAATAGCGTGTGTTATGATCATACTATGATAGGTAGTGGCGCTATTATTCTCC 13500

F V T I S T N T H I G R F F H A N I Y S

TTTGTTACAATATCTACTAATACTCATATAGGGAGGTTTTCATGCAACATATACTCA 13560

Y V A H D C Q I G D Y V T F A P G A K K C

TACGTTGCACATGATTGTCAAAATAGGAGACTATGTACATTGCTCCTGGGCTAAATGT 13620

N G Y V V I E D N A Y I G S G A V I K Q

AATGGATATGTTGTTATGAAGACAATGCATATATAGGCTCGGGTGCAATTAAGCAG 13680

G V P N R P L I I G A G A I I G M G A V

GGTGTTCTTAATCGCCCACTTATTATGGCGCGGGGACATTATAGGTATGGGGGCTGTT 13740

V T K S V P A G I T V C G N P A R E M K

GTCACAAAAGTGTCTCGCGGTATAACTGTGTGCGGAAATCCAGCAAGAGAAATGAAA 13800

**End of orf12**

R S P T S I \*

AGATCGCCAACATCTATTTAATGGGAATGCGAAAACACGTTCCAAATGGGACTAATGTTT 13860

WO 99/61458

PCT/AU99/00385

28/96

AAAATATATATAATTCGCTAATTTACTAAATTATGGCTTCTTTTAAGCTATCCTTTAC 13920

TTAGTTATTACTGATACAGCATGAAATTTATAATACTCTGATACATTTTATACGTTATT 13980

CAAGCCGCATATCTAGCGGTAACCCCTGACAGGAGTAAACAATG 14024

13920  
13980  
14024  
14080  
14140  
14200  
14260  
14320  
14380  
14440  
14500  
14560  
14620  
14680  
14740  
14800  
14860  
14920  
14980  
15040  
15100  
15160  
15220  
15280  
15340  
15400  
15460  
15520  
15580  
15640  
15700  
15760  
15820  
15880  
15940  
16000  
16060  
16120  
16180  
16240  
16300  
16360  
16420  
16480  
16540  
16600  
16660  
16720  
16780  
16840  
16900  
16960  
17020  
17080  
17140  
17200  
17260  
17320  
17380  
17440  
17500  
17560  
17620  
17680  
17740  
17800  
17860  
17920  
17980  
18040  
18100  
18160  
18220  
18280  
18340  
18400  
18460  
18520  
18580  
18640  
18700  
18760  
18820  
18880  
18940  
19000  
19060  
19120  
19180  
19240  
19300  
19360  
19420  
19480  
19540  
19600  
19660  
19720  
19780  
19840  
19900  
19960  
20020  
20080  
20140  
20200  
20260  
20320  
20380  
20440  
20500  
20560  
20620  
20680  
20740  
20800  
20860  
20920  
20980  
21040  
21100  
21160  
21220  
21280  
21340  
21400  
21460  
21520  
21580  
21640  
21700  
21760  
21820  
21880  
21940  
22000  
22060  
22120  
22180  
22240  
22300  
22360  
22420  
22480  
22540  
22600  
22660  
22720  
22780  
22840  
22900  
22960  
23020  
23080  
23140  
23200  
23260  
23320  
23380  
23440  
23500  
23560  
23620  
23680  
23740  
23800  
23860  
23920  
23980  
24040  
24100  
24160  
24220  
24280  
24340  
24400  
24460  
24520  
24580  
24640  
24700  
24760  
24820  
24880  
24940  
25000  
25060  
25120  
25180  
25240  
25300  
25360  
25420  
25480  
25540  
25600  
25660  
25720  
25780  
25840  
25900  
25960  
26020  
26080  
26140  
26200  
26260  
26320  
26380  
26440  
26500  
26560  
26620  
26680  
26740  
26800  
26860  
26920  
26980  
27040  
27100  
27160  
27220  
27280  
27340  
27400  
27460  
27520  
27580  
27640  
27700  
27760  
27820  
27880  
27940  
28000  
28060  
28120  
28180  
28240  
28300  
28360  
28420  
28480  
28540  
28600  
28660  
28720  
28780  
28840  
28900  
28960  
29020  
29080  
29140  
29200  
29260  
29320  
29380  
29440  
29500  
29560  
29620  
29680  
29740  
29800  
29860  
29920  
29980  
30040  
30100  
30160  
30220  
30280  
30340  
30400  
30460  
30520  
30580  
30640  
30700  
30760  
30820  
30880  
30940  
31000  
31060  
31120  
31180  
31240  
31300  
31360  
31420  
31480  
31540  
31600  
31660  
31720  
31780  
31840  
31900  
31960  
32020  
32080  
32140  
32200  
32260  
32320  
32380  
32440  
32500  
32560  
32620  
32680  
32740  
32800  
32860  
32920  
32980  
33040  
33100  
33160  
33220  
33280  
33340  
33400  
33460  
33520  
33580  
33640  
33700  
33760  
33820  
33880  
33940  
34000  
34060  
34120  
34180  
34240  
34300  
34360  
34420  
34480  
34540  
34600  
34660  
34720  
34780  
34840  
34900  
34960  
35020  
35080  
35140  
35200  
35260  
35320  
35380  
35440  
35500  
35560  
35620  
35680  
35740  
35800  
35860  
35920  
35980  
36040  
36100  
36160  
36220  
36280  
36340  
36400  
36460  
36520  
36580  
36640  
36700  
36760  
36820  
36880  
36940  
37000  
37060  
37120  
37180  
37240  
37300  
37360  
37420  
37480  
37540  
37600  
37660  
37720  
37780  
37840  
37900  
37960  
38020  
38080  
38140  
38200  
38260  
38320  
38380  
38440  
38500  
38560  
38620  
38680  
38740  
38800  
38860  
38920  
38980  
39040  
39100  
39160  
39220  
39280  
39340  
39400  
39460  
39520  
39580  
39640  
39700  
39760  
39820  
39880  
39940  
40000  
40060  
40120  
40180  
40240  
40300  
40360  
40420  
40480  
40540  
40600  
40660  
40720  
40780  
40840  
40900  
40960  
41020  
41080  
41140  
41200  
41260  
41320  
41380  
41440  
41500  
41560  
41620  
41680  
41740  
41800  
41860  
41920  
41980  
42040  
42100  
42160  
42220  
42280  
42340  
42400  
42460  
42520  
42580  
42640  
42700  
42760  
42820  
42880  
42940  
43000  
43060  
43120  
43180  
43240  
43300  
43360  
43420  
43480  
43540  
43600  
43660  
43720  
43780  
43840  
43900  
43960  
44020  
44080  
44140  
44200  
44260  
44320  
44380  
44440  
44500  
44560  
44620  
44680  
44740  
44800  
44860  
44920  
44980  
45040  
45100  
45160  
45220  
45280  
45340  
45400  
45460  
45520  
45580  
45640  
45700  
45760  
45820  
45880  
45940  
46000  
46060  
46120  
46180  
46240  
46300  
46360  
46420  
46480  
46540  
46600  
46660  
46720  
46780  
46840  
46900  
46960  
47020  
47080  
47140  
47200  
47260  
47320  
47380  
47440  
47500  
47560  
47620  
47680  
47740  
47800  
47860  
47920  
47980  
48040  
48100  
48160  
48220  
48280  
48340  
48400  
48460  
48520  
48580  
48640  
48700  
48760  
48820  
48880  
48940  
49000  
49060  
49120  
49180  
49240  
49300  
49360  
49420  
49480  
49540  
49600  
49660  
49720  
49780  
49840  
49900  
49960  
50020  
50080  
50140  
50200  
50260  
50320  
50380  
50440  
50500  
50560  
50620  
50680  
50740  
50800  
50860  
50920  
50980  
51040  
51100  
51160  
51220  
51280  
51340  
51400  
51460  
51520  
51580  
51640  
51700  
51760  
51820  
51880  
51940  
52000  
52060  
52120  
52180  
52240  
52300  
52360  
52420  
52480  
52540  
52600  
52660  
52720  
52780  
52840  
52900  
52960  
53020  
53080  
53140  
53200  
53260  
53320  
53380  
53440  
53500  
53560  
53620  
53680  
53740  
53800  
53860  
53920  
53980  
54040  
54100  
54160  
54220  
54280  
54340  
54400  
54460  
54520  
54580  
54640  
54700  
54760  
54820  
54880  
54940  
55000  
55060  
55120  
55180  
55240  
55300  
55360  
55420  
55480  
55540  
55600  
55660  
55720  
55780  
55840  
55900  
55960  
56020  
56080  
56140  
56200  
56260  
56320  
56380  
56440  
56500  
56560  
56620  
56680  
56740  
56800  
56860  
56920  
56980  
57040  
57100  
57160  
57220  
57280  
57340  
57400  
57460  
57520  
57580  
57640  
57700  
57760  
57820  
57880  
57940  
58000  
58060  
58120  
58180  
58240  
58300  
58360  
58420  
58480  
58540  
58600  
58660  
58720  
58780  
58840  
58900  
58960  
59020  
59080  
59140  
59200  
59260  
59320  
59380  
59440  
59500  
59560  
59620  
59680  
59740  
59800  
59860  
59920  
59980  
60040  
60100  
60160  
60220  
60280  
60340  
60400  
60460  
60520  
60580  
60640  
60700  
60760  
60820  
60880  
60940  
61000  
61060  
61120  
61180  
61240  
61300  
61360  
61420  
61480  
61540  
61600  
61660  
61720  
61780  
61840  
61900  
61960  
62020  
62080  
62140  
62200  
62260  
62320  
62380  
62440  
62500  
62560  
62620  
62680  
62740  
62800  
62860  
62920  
62980  
63040  
63100  
63160  
63220  
63280  
63340  
63400  
63460  
63520  
63580  
63640  
63700  
63760  
63820  
63880  
63940  
64000  
64060  
64120  
64180  
64240  
64300  
64360  
64420  
64480  
64540  
64600  
64660  
64720  
64780  
64840  
64900  
64960  
65020  
65080  
65140  
65200  
65260  
65320  
65380  
65440  
65500  
65560  
65620  
65680  
65740  
65800  
65860  
65920  
65980  
66040  
66100  
66160  
66220  
66280  
66340  
66400  
66460  
66520  
66580  
66640  
66700  
66760  
66820  
66880  
66940  
67000  
67060  
67120  
67180  
67240  
67300  
67360  
67420  
67480  
67540  
67600  
67660  
67720  
67780  
67840  
67900  
67960  
68020  
68080  
68140  
68200  
68260  
68320  
68380  
68440  
68500  
68560  
68620  
68680  
68740  
68800  
68860  
68920  
68980  
69040  
69100  
69160  
69220  
69280  
69340  
69400  
69460  
69520  
69580  
69640  
69700  
69760  
69820  
69880  
69940  
70000  
70060  
70120  
70180  
70240  
70300  
70360  
70420  
70480  
70540  
70600  
70660  
70720  
70780  
70840  
70900  
70960  
71020  
71080  
71140  
71200  
71260  
71320  
71380  
71440  
71500  
71560  
71620  
71680  
71740  
71800  
71860  
71920  
71980  
72040  
72100  
72160  
72220  
72280  
72340  
72400  
72460  
72520  
72580  
72640  
72700  
72760  
72820  
72880  
72940  
73000  
73060  
73120  
73180  
73240  
73300  
73360  
73420  
73480  
73540  
73600  
73660  
73720  
73780  
73840  
73900  
73960  
74020  
74080  
74140  
74200  
74260  
74320  
74380  
74440  
74500  
74560  
74620  
74680  
74740  
74800  
74860  
74920  
74980  
75040  
75100  
75160  
75220  
75280  
75340  
75400  
75460  
75520  
75580  
75640  
75700  
75760  
75820  
75880  
75940  
76000  
76060  
76120  
76180  
76240  
76300  
76360  
76420  
76480  
76540  
76600  
76660  
76720  
76780  
76840  
76900  
76960  
77020  
77080  
77140  
77200  
77260  
77320  
77380  
77440  
77500  
77560  
77620  
77680  
77740  
77800  
77860  
77920  
77980  
78040  
78100  
78160  
78220  
78280  
78340  
78400  
78460  
78520  
78580  
78640  
78700  
78760  
78820  
78880  
78940  
79000  
79060  
79120  
79180  
79240  
79300  
79360  
79420  
79480  
79540  
79600  
79660  
79720  
79780  
79840  
79900  
79960  
80020  
80080  
80140  
80200  
80260  
80320  
80380  
80440  
80500  
80560  
80620  
80680  
80740  
80800  
80860  
80920  
80980  
81040  
81100  
81160  
81220  
81280  
81340  
81400  
81460  
81520  
81580  
81640  
81700  
81760  
81820  
81880  
81940  
82000  
82060  
82120  
82180  
82240  
82300  
82360  
82420  
82480  
82540  
82600  
82660  
82720  
82780  
82840  
82900  
82960  
83020  
83080  
83140  
83200  
83260  
83320  
83380  
83440  
83500  
83560  
83620  
83680  
83740  
83800  
83860  
83920  
83980  
84040  
84100  
84160  
84220  
84280  
84340  
84400  
84460  
84520  
84580  
84640  
84700  
84760  
84820  
84880  
84940  
85000  
85060  
85120  
85180  
85240  
85300  
85360  
85420  
85480  
85540  
85600  
85660  
85720  
85780  
85840  
85900  
85960  
86020  
86080  
86140  
86200  
86260  
86320  
86380  
86440  
86500  
86560  
86620  
86680  
86740  
86800  
86860  
86920  
86980  
87040  
87100  
87160  
87220  
87280  
87340  
87400  
87460  
87520  
87580  
87640  
87700  
87760  
87820  
87880  
87940  
88000  
88060  
88120  
88180  
88240  
88300  
88360  
88420  
88480  
88540  
88600  
88660  
88720  
88780  
88840  
88900  
88960  
89020  
89080  
89140  
89200  
89260  
89320  
89380  
89440  
89500  
89560  
89620  
89680  
89740  
89800  
89860  
89920  
89980  
90040  
90100  
90160  
90220  
90280  
90340  
90400  
90460  
90520  
90580  
90640  
90700  
90760  
90820  
90880  
90940  
91000  
91060  
91120  
91180  
91240  
91300  
91360  
91420  
91480  
91540  
91600  
91660  
91720  
91780  
91840  
91900  
91960  
92020  
92080  
92140  
92200  
92260  
92320  
92380  
92440  
92500  
92560  
92620  
92680  
92740  
92800  
92860  
92920  
92980  
93040  
93100  
93160  
93220  
93280  
93340  
93400  
93460  
93520  
93580  
93640  
93700  
93760  
93820  
93880  
93940  
94000  
94060  
94120  
94180  
94240  
94300  
94360  
94420  
94480  
94540  
94600  
94660  
94720  
94780  
94840  
94900  
94960  
95020  
95080  
95140  
95200  
95260  
95320  
95380  
95440  
95500  
95560  
95620  
95680  
95740  
95800  
95860  
95920  
95980  
96040  
96100  
96160  
96220  
96280  
96340  
96400  
96460  
96520  
96580  
96640  
96700  
96760  
96820  
96880  
96940  
97000  
97060  
97120  
97180  
97240  
97300  
97360  
97420  
97480  
97540  
97600  
97660  
97720  
97780  
97840  
97900  
97960  
98020  
98080  
98140  
98200  
98260  
98320  
98380  
98440  
98500  
98560  
98620  
98680  
98740  
98800  
98860  
98920  
98980  
99040  
99100  
99160  
99220  
99280  
99340  
99400  
99460  
99520  
99580  
99640  
99700  
99760  
99820  
99880  
99940  
100000

29/96

ATGGCACAGTCATTAAATACCAACAGCCCTCTCGCTGATCACTCAAAATAATATCAACAAG  
 AACCAAGCTCTGCGCTGTTCGAGTTCTATCGAGCGTCTGTCTTCTGGCTTGCGTATTAAACAGC  
 GCGAAGGATGACGCCGCCGGGTGAGGCGATTGCTAACCGTTTACTTCTAACATTAAAGGC  
 CTGACTCAGGCTGCACGTAAACGCCAACGACGGTATTCTGTGTGCACAGACCACTGAAGGC  
 GCGCTGTCCGAAATCAACAACAACCTTACAGCGTATCCGTGAGCTGACGGTTCAGGCTTCT  
 ACCGGGACTAACTCTGATTCCGATCTGGACTCCATTTCAGGACGAAATCAATCCCGTCTC  
 GACGAAATTGACCGCGGTATCCGGTCAGACCCAGTTCACCGCGTGAACGCTACTGGCAAAA  
 GACGGTTCGATGAAATTCAGGTAGTGCAGAACGACGCCAGACTATCACTATTGATCTG  
 AAGAAATTGACTCTGATACGCTGGGCTGAATGGTTTAACTGAATGGTTCCGGTACG  
 ATAGCCAATAAAGCGGCGACCAATTAGCGACCTGACAGCAGCGAAATGGATGCTGCAACT  
 AATACTATAACTACAACAAATAATGCGCTGACTGCATCAAAGGCCCTTGATCAACTGAAA  
 GATGGTGACACTGTTACTATCAAAGCAGATGCAGCTCAAACCTGCCACGGTCTATACATAC  
 AATGCATCTGCTGGTAACTTCTCATTTCAGTAATGTATCGAATAATACTTCAGCAAAAGCA  
 GGTGATGTAGCAGCTAGCCTTCTCCCGCGGCTGGGCAAACTGCTAGTGGTGTTCACAAA  
 GCAGCAAGCGGTGAAGTGAACCTTTGATGTTGATGCGAATGGTAAATTAACAATCGGAGGA  
 CAGGAAGCCCTATTTAAGTGTGATGGTAACTTAACTACAAACGATGCTGGTGGTGCAGCT  
 GCGGCTACGCTTGATGGTTTATTCAAGAAAGCTGGTGTGATGCTCAATCAATCGGGTTTAAAT  
 AAGACTGCATCAGTCACGATGGGGGAACAACCTTATAACTTTAAACCGGGTCTGATGCT  
 GGTGCTGCAACTGCTAACGCGGGGTATCGTTCACTGATACAGCTAGCAAGAAACCGTT  
 TTAATAAAGTGGCTACAGCTAAACAAGGCACAGCAGTTGCGAGCTAACGGTGATACATCC  
 GCAACAATTACCTATAAATCTGGCGTTCAGACGTCAGGCGGTATTGCGCGCAGGTGAC  
 GGTACTGCTAGCGCAAAATATCCGCGATACTGACGTTTCTAATGCAACAGCAACATAC  
 ACAGATGCTGATGGTGAATGACTACAATTGGTTTATACACCACGAAGTATTCAATCGAT  
 GCTAACACCGCAAGGTAACCTGTTGATTCTGGAACGTGGTTCGGGTAAATATGCGCGGAAA  
 GTGGGGCTGAAGTATATGTTAGTGCTAATGGTACTTTAACACAGATGCAACTAGCGAA  
 GGCACAGTAACAAAAGATCCACTGAAAGCTCTGGATGAAGCTATCAGCTCCATCGACAAA  
 TTCCGTTTATCCCTGGGGCTATCCAAAACCGTTTGGATTCCGCGCTACCAACCTGAAC  
 AACCACTACCAACCTGCTGAAAGCGAGTCCCGTATTTCAGGACGCCGACTATGCGACC  
 GAAGTGTCCAACATGTCGAAAGCGCAGATTATCCAGCAGGCGGTAACTCCGTGCTGGCA  
 AAAGCCAACAGGATCCGCGAGCAGGTTCTGTCTCTACTGCGAGGTTAA

Figure 7

30/96

AACAAATCTCAGTCTTCTCTTAGCTCTGCTATT  
 GAGCGTCTGTCTTCTGGTCTGCGTATTACAGCGCAAAAGACGATGCAGCAGGTGAGCGG  
 ATTGCTAACCGTTTTACGGCAAAATATTAAAGGTCTGACCCAGGCTTCCCGTAAACGCAAAAT  
 GATGGTATTCTGTGCGCAGACCACTGAAGGTGCGCTGAATGAAATTAACAACAACCTG  
 CAGCGTATTCTGTAACCTTCTGTTCAGGCAACTAACGGTACTAACTCTGACAGTGACCTG  
 ACCTCCATCCAGTCCGAAATCCAGCAGCGTCTGAGTGAAATTGACCGTGTTCCTGGTCTAG  
 ACTCAGTTTAAACGGCGTTAAAGTGTGGCTTCTGATCAGGATATGACTATTACAGTTGGT  
 GCAACGACGCGGAAACAATTACTATTAACTGCAGGAAATTAATCCGACACACTGGGA  
 TTATCTGGTTTTGGTATTAAAGATCCTACTAAATTAAAGCCGCAACGGCTGAAACAACC  
 TATTTGGATCGACAGTTAAGCTTGTGACGCTAATACACTTGATGAGATATTACAGCT  
 ACAGTTAAAGGCACTACGACTCCGGGCCAACGTGACGCTAATATTATGTCTGATGCTAAC  
 GGTAAGTTGTACGTTAAAGTTCCGGTTCAGATAAACCCGCTGAAAATGGTTATTATGAA  
 GTTACTGTGGAGGATGATCCGACATCTCCTGATGCAGGTAAGCTGAAGCTGGGGGCTCTA  
 GCGGGTACCCAGCCTCAAGCTGGTAATTTAAAGGAAGTCACAACGGTGAAAGGGAAGGGG  
 GCTATTGATGTTTCACTGGGTACTGATACCGCAACCGCTTCTATCACAGGTGCAAAACTC  
 TTTAAGTTAGAAGACGCCAATGGCAAAGATACTGGTTCAATTTGCGTTGATTGGTGATGAC  
 GGTAAACAGTATGCAGCGAATGTTGATCAGAAAACAGGAGCAGTTTCGGTTAAACAATG  
 TCTTACACTGATGCTGACGGTGTCAAACACGACAATGTTAAAGTTGAACTGGGTGGAAGC  
 GATGGCAAAACCGAAGTTGTAACGTCAACCGATGGCAAACTTACAGTGTTAGTGATTTA  
 CAAGGTAAGAGCCTGAAAACGATTCTATTGCAGCAATTTCTACGCAAAAACAGAAGAT  
 CCTTGGGTGCTATCGATAAAGCACTGTCTCAGGTTGACTCGTTGCGTTCTAACCTAGGT  
 GCAATTCAAAATCGTTTCGACTCTGCCATCACCACCTTGGCAACACCGTAAACAACCTG  
 TCTTCTGCCGTAGCCGTATCGAAGATGCTGACTACGCGACCGAAGTGTCTAACATGTCT  
 CGTGGCAGATCCTGCAACAACGGGTACCTCTCTTCTGGCGCAG

Figure 8

31/96

AACAAATCTCAGTCTTCTCTGAGCTCCGCCATTGAACGTCCTCTTC  
 TGGCCTGCGTATTAAACAGTGCTAAAGATGACGACGAGGTGAGGCGATTGCTAACCGTTT  
 TACAGCAAATATTAAAGTCTGACTCAGGCTTCCCGTAACGCGAATGATGGTATTCTGT  
 TGCGCAGACCACTGAAGGTGCGCTTCTGAAATCAACAATAACTTACAGCGTATTCTGTGA  
 ATTGTGAGTACAGGCCACTAATGGTACAACTCTGACTCCGACCTGAATTC AATT CAGGA  
 TGAATTTACACAACGCCCTTAGTGAAATTGATCGTGTTTCTAACGACACAAATTTAATGG  
 TGTAAAGTTCTGGCTTCTGATCAGACTATGAAATTC AAGTAGGTGCGAACGATGGTGA  
 AACCATTGAGATTGCCCTTGATAAAATGATGCTAAAACTTGGGGCTTGATAACTTTAG  
 CGTAGCACCAGGAAAAGTTCCAATGTCCTCTGCGGTTGCACCTTAAGAGCGAAGCCGCTCC  
 TGACTTAAC TAAGGTAATGCAACTGATGGTAGTGTGGGAGGTGCTAAAGCATTCCGSTAT  
 CAATTTATAAAAATGCTGATGTTGAAACTTATTTTGGTACCGGTAATGTACAAGATACAAA  
 GGATACAACTGATGCGACCGGTACTG CAGGAACAAAAGTTTATCAAGTACAGGTGGAAGG  
 GCAGACTTATTTTGTGGTCAAGATAATAATACCAACACGAACGGTTTACATTATTGAA  
 ACAAACTCTACAGGTTATGAAAAAGTT CAGGTGGGTGGTAAGGATGTT CAGTTAGCAAA  
 CTTTGGTGGTGGTGAATGCTGATTTGTTGAAGATAATGGTTCTGCCACATCAGTTGATTT  
 AGCTGCGGGTAAAAAGGTTAAAGCATTAGCTTATAATGATGACCAATGTCTGTTTATTT  
 TGGGGGAAAAAACCTAGATGTCCACCAAGTACAAGATACCCAAGGGAATCCTGTACCTAA  
 TTCATTTGCTGCTAAAAACATCAGACGGCACCTACATTG CAGTAATGTAGATGCCGCTAC  
 AGGTAACACGTCCTGTTATTACTGATCCTAATGCTAAGGCAGTTGAATGGGCAGTAAAAAA  
 TGATGGTTCTGCACAGGCAATTATGCGTGAAGATGATAAGGTTTATACAGCCAATATCAC  
 GAATAAGACGGCAACCAAAGGTGCTGAACTCAGTGCCCTCAGATTGAAAGCCTTAGCAAC  
 CACAAATCCATTATCCACATTAGACGAAGCTTTGGCAAAAGTTGATAAGTTGCGCAGTTC  
 TTTGGGTG CAGTACAAAACCGTTTCGACTCTGCCATCAACCACTTGGCAACACCGTAA  
 CAACCTGTCTTCTGCCCCGTAGCCGTATAGAAGATGCTGACTACGCAACCGAAGTGTCTAA  
 CATGTCTCGTGCGCAGATCCTGCAACCAAGCGGTAACCTGTGTTCTGSCACAG

Figure 9

32/96

AACAAAAACCAAGTCTGCGCTGTCGACTTCTATCGAG  
 CGCCTCTCTTCTGGTCTGCGTATTAACAGCGCTAAAGATGACGCCGCGGGCCAGGCGATT  
 GCTAACCGCTTTACTTCTAACATCAAAGGTCTGACTCAGGCCGCACGTAACGCCAACGAC  
 GGTATTTCTCTGGCGCAGACGGCTGAAGGCGCGCTGTGACAGATTAAACAACAATTGCAG  
 CGTATTCGTGAAGTACCGTTTACGGCTCTACCGGCACGAACCTGATTCCGACCTGTCT  
 TCTATTGAGGACGAAATCAAATCCCGTCTTGATGAAATTGACCGTGTATCTGGTCAGACC  
 CAGTTCAACGGTGTGAACGTGCTGTGAAAAACGATTGATGAAGATTGAGATTGGTGCC  
 AATGATAACCGACGATCAGCATTTGGCTTGCAACAAATCGACAGTACCACTTTGAATCTG  
 AAAGGATTTACCGTGTCCGGCATGGCGGATTTGACGCGCGGCAAACTGACGGCTGCTGAT  
 GGTACAGCAATTGCTGCTGCGGATGTCAAGGATGCTGGGGGTAAACAAGTCAATTTACTG  
 TCTTACACTGACACCGCGTCTAACAGTACTAAATATCGGTGCGTTGATTCTGCAACCGGT  
 AAATACATGGAAGCCACTGTAGTCATTACCGGTACGGCGCGCGGTAACTGTTGGTGCA  
 GCGGAAGTGCGCGGAGCGGTACAGCCGATCCGTTAAAAAGCACTGGGATGCGCAATCGCT  
 AAAGTCGACAAATTCGCTCTCTCCCTCGGTGCCGTTCAAAACCGTCTGGAATTCGCGGT  
 ACCAACCTGAACAACACCAACCAACCTGTCTGAAGCGAGTCCCGTATTGAGGACGCC  
 GACTATGCGACCGAAGTGTCCAACATGTCGAAAGCGCAGATTATCCAGCAGGCGGGCAAC TCCGTGCTGTCTAA

Figure 10

33/96

AACAAAAACAGTCTGCGCTGTCGACTTCTAT  
 CGAGCGCCTCTCTTCTGGTCTGCGTATTAACAGCGCTAAAGATGACGCCGCGGGCCAGGC  
 GATTGCTAACCGCTTCACTTCTAACATCAAAGGTCTGACTCAGGCCGCACGTAACGCCAA  
 CGACGGTATCTCTGCGCGCAGACCACTGAAGGCGCGCTGCTCTGAAATCAACAACAACTT  
 GCAGCGTGTGCGTGAGTTGACCGTTCAGGCGACGACCGGGACTAACTCTGATTCTGACCT  
 GTCTTCTATTCAAGACGAAATCAAATCCCGTCTGGATGAAATGATCGCGTTTCCGGTCA  
 GACCCAGTTCAACGGCGTGAATGTGCTGGCGAAAGATGTTTCGATGAAGATTCAAGTTGG  
 CGCGAATGATGGGCAGACTATTAGCATTGATTGCGAAGATTGACTCTTCTACATTAGG  
 ACTGAACGGTTTCTCCGTTTCGGGTCAGTCACTTAACGTTAGTGATTCCATTACTCAAAT  
 TACCGGTGCGCCGGGACAAAACCTGTTGGTGTGATTCTCACTGCTGTTGCGAAAGATCT  
 GACTACTGCGACAGGTAAAACAGTCGATGTTTCTAGCCTGACGTTACACAACACTCTGGA  
 TGCGAAAGGGGTGCTACATCACAGTTCGTGTTCAATCCGGCAATGATTTCTACTCCGC  
 GTCGATTAAATCATACAGACGGCAAAGTCAAGTTGAATAAAGCCGATGTCGAATACACAGA  
 CACCGATAATGGACTAACGACTGCGGCTACTCAGAAAGATCAACTGATTAAAGTTGCCGC  
 TGACTCTGACGGCTCGGCTGCGGGATATGTAACATTCGAAGTAAAACTACGCTACAAAC  
 GGTTTCAACGGCACTTGATGATAATACTGCGGCAAAAGCAACAGATAATAAAGTTGTTGT  
 TGAATTATCAACAGCAAAACCGACTGCACAGTTCTCAGGGGCTTCTTCTGCTGATCCACT  
 GGCACTTTATAGACAAAGCTATTGCACAGGTTGATACTTTCCGCTCCTCCCTCGGTGCGGT  
 GCAAAACCGTCTGGATTCCGCAGTAACCAACCTGAACAACACCAACCAACCTGTCTGA  
 AGCGCAGTCCCGTATTCAAGGACGCCGACTATGCTACAGAAGTGTCACACATGTCGAAAGC  
 GCAGATCATCCAGCAGGCAGGTAACTCGGTGCTGTCCAAA

Figure 11



34/96

ATGGCACAAAGTCATTAATACCAACAGCCTCTCGC  
 TGATCACTCAAAATAATATCAACAAGAACAGTCTGCGCTGTCGAGTTCTATCGAGCGTC  
 TGTCTTCTGGCTTGCGTATTAAACAGCGCGAAGGATGACGCCGCGGTCAGGCGATTGCTA  
 ACCGTTTTACTTCTAACATTAAAGGCCCTGACTCAGGCTGCACGTAAACCAACGACGGTA  
 TTTCTGTGCGCAGACCAACCGAAGGCGCGCTGTCCGAAATTAAACAACACTTACAGCGTA  
 TTCGTGAAGTACGCGTTACGGCTTCTACCGGGACTAACTCTGATTCTGGATCTGGACTCCA  
 TTCAGGACGAAATCAAATCCGCTCTGCACGAAATTGACCGCGTATCCGGTCAGACCCAGT  
 TCAACGGCGTGAACGTACTGGCAAAAGACGGTTCGATGAAAATTACGGTTGGTGCGAATG  
 ACGGCCAGACTATCACTATTGATCTGAAGAAAATTGACTCTGTATACGCTGGGGCTGAATG  
 GGTTTAATGTGAACGGCAAGGGGAAACGGCTAATACGGCAGCAACCTGAAAGATATGT  
 CTGGATTACAGCTGCGGCGGCACCGGGGAACTGTTGGGTAACTCAATATACTGACA  
 AATCGGCTGTAGCAAGTAGCGTAGATATTCTAAATGCTGTTGCTGGCGCAGATGGAAATA  
 AAGTTACAAC TAGCGCCGATGTTGGTTTTGGTACACGAGCGCTGCTGTAACCTATACCT  
 ACAATAAAGACACTAATTCATATTCGCGCGCTTCTGATGATATTTCCAGCGCTAACCTGG  
 CTGCTTTCCCTCAATCCTCAGGCCGGAGATAACGACTAAAGCTACAGTTACAATTGGTGGCA  
 AAGATCAAGATGTAACATCGATAAAATCCGGTAATTTAACTGCTGCTGATGATGGCGCAG  
 TACTTTATATGGATGCTACCGGTAACTTAACATAAAATAATGCTGGTGGTGATACACAAG  
 CTACTTTGGCTAAACTTGCTACTGCTACTGGTGCTAAAGCGCGACCATCCAAACTGATA  
 AAGGAACATTCACCAGTGACCGGTACAGCGTTTGATGGTGATCAATGTCCATTGATACCA  
 ATACATTTGCAANTGCAGTAAAAAATGACACTTATCTGCCACTGTAGTGTCTAAGACTT  
 ATAGCGTAACAACAGGTTCTGCTGCTGCAGACACCGCTTATATGAGCAATGGGGTTCTCA  
 GTGATACTCCGCCAACTTACTATGCAACAAGCTGATGGAAGTATCACAACTACTGAGGATG  
 CGGCTGCCGGTAAACTGGTCTACAAAGGTTCCGATGGTAAGTTAAACAACGGATACGACTA  
 GCAAAGCAGAATCAACATCAGATCCGCTGGCAGCTCTTGACGACGCTATCAGCCAGATCG  
 ACAAAATCCGCTCCTCCTGGGTGCGGTGCAAAACCGTCTGGATTCCGCGAGTGACCAACC  
 TGAACAACACCACTACCAACCTGTCTGAAGCGCAGTCCCGTATTACGAGCGCGGACTATG  
 CGACCGAAGTGTCCAACATGTGAAAGCGCAGATTATCCAGCAGCGCGTAACTCCGTGC  
 TGGCAAAAGCTAACCAAGTTCCGCAGCAGGTTCTGTCTCTGCTGCAGGGTTAA

Figure 12

35/96

ATGGCACAAG TCATTAATAC CAACAGCCTC TCGCTGATCA CTCAAAATAA TATCAACAAG  
 AACCAAGTCTG CGCTGTCGAG TCTATCGAG CGTCTGTCTT CTGGCTTGCG TATTAAACAGC  
 GCGAAGGATG ACGCCGCGGG TCAGGCGATT GCTAACCGTT TTACTTCTAA CATTAAAGGC  
 CTGACTCAGG CTGCACGTAA CGCCAAACGAC GGTATTCTTG TTGCACAGAC CACCGAAGGC  
 GCGCTGTCTG AAATCAACAA CAACTTACAG CGTATCCGTG AGCTGACGGT TCAGGCTTCT  
 ACCGGAACCT ACTCTGATTG GGATCTGGAC TCCATTTCAGG ACGAAATCAA ATCCCGTCTT  
 GATGAAATTG ACCGCGTATC CGGCCAGACC CAGTTCAACG CGTGAAACGT ACTGGCAAAA  
 GACGGTTCGA TGAATAATCA GGTGGGTGCG AATGACGGTG AAATATCAC TATCGACCTG  
 AAGAAAATCG ATTCTGATAC TCTGGGTCTG AATGGTTTTA ACGTAAATGG TAAAGGTACT  
 ATTACCAACA AAGCTGCAAC GGTAAAGTGAT TTAACCTCTG CTGGCGCGAA GTTAAACAC  
 CACGACAGGT CTTTATGATC TGAAAACCGA AAATACCTTG TTAACCTACC ATGCTGCATT  
 CGATAAATTA GGAATGGCG ATAAAGTCAC CGTTGGCGGC GTAGATTATA CTTACAAACGC  
 TAAATCTGGT GATTTTACTA CCACCAATC TACTGCTGGT ACGGGTGTAG ACGCCGCGGC  
 GCAGGCTACT GATTGAGCTA AAAAACGTGA TGCGTTAGTG GCCACCTTTC ATGCTGATGT  
 GGGTAAATCT GTTAATGGTT CTTACACCAC AAAAGATGGT ACTGTTTCTT TCGAAACCGA  
 TTCAGCAGGT AATATCACCA TCGGTGGAAG CCAGGCATAC GTAGACGATG CAGGCAACTT  
 GACGACTAAC AACGCTGGTA GCGCAGCTAA AGCTGATATG AAAGCGCTGC TTAAGCCGCG  
 GAGCGAAGGT AGTGACGGTG CTTCTCTGAC ATTCAATGGC ACTGAATATA CTATCGCAAA  
 AGCAACTCCT GCGACAACCT CTCAGTAGC TCCGTTAATC CCTGGTGGGA TTACTTATCA  
 GGCTACAGTG AGTAAAGATG TAGTATTGAG CGAAACCAAA CGCGCTGCCG CGACATCTTC  
 AATTACCTTT AATTCCGGTG TACTGAGCAA AACTATTGGG TTACCGCGG GTGAATCCAG  
 TGATGCTGCG AAGTCTTATG TGGATGATAA AGGTGGTATT ACTAACGTTG CCGACTATAC  
 AGTCTCTTAC AGCGTTAACA AGGATAACGG CTCTGTGACT GTTGCCGGGT ATGCTTCAGC  
 GACTGATACC AATAAAGATT ATGCTCCAGC AATTGGTACT GCTGTAATG TGAACCTCGC  
 GGGTAAATC ACTACTGAGA CTACCAAGTC TGGTTCTGCA ACGACCAACC CGCTTGCTGC  
 CCTGGACGAC GCTATCAGCT CCATCGACAA ATTCCGTTCT TCCTGGGTG CTATCCAGAA  
 CCGTCTGGAT TCCGAGTCA CCAACCTGAA CAACACCACT ACCAACCTGT CTGAAGCGCA  
 GTCCCGTATT CAGGACGCCG ACTATGCGAC CGAAGTGTC AACATGTCGA AAGCGCAGAT  
 TATCCAGCAG GCCGGTAAC TCGTGTGGC AAAAGCCAAC CAGGTACCGC AGCAGGTTCT  
 GTCTCTGCTG CAGGGTTAA

Figure 13

36/96

AACAAATCTCAGTCTTCTCTTAGCTCTGCTA  
 TTGAGCGTCTGTCTTCTGGTCTGCGTATTAACAGCGCAAAAGACGATGACGAGGTCCAGG  
 CGATTGCTAACCGTTTTACGGCAAATATTAAAGGTCTGACCCAGGCTTCCCCTAACGCAA  
 ATGATGGTATTTCTGTTGCGCAGACCACTGAAGGTGCGCTGAATGAAATTAACAACAACC  
 TGCAGCGTATTCGTGAACCTTCTGTTTCAGGCAACTAACGGTACTAACTCTGACAGCGATC  
 TTTCTTCTATCCAGGCTGAAATTTACTCAACGCTCTGGAAGAAATTGACCGGTATCTGAGC  
 AAACCTCAGTTTAACGGCGTGAAAGTCCTTGCTGAAAATAATGAAATGAAATTCAGGTG  
 GTGCTAATGATGGTGAAACCATCACTATCAATCTGGCAAAAATTGATGCGAAAACTCTCG  
 GCCTGGACGGTTTTAATATCGATGGCGCGCAGAAAGCAACAGGCAGTGACCTGATTCTTA  
 AATTTAAAGCGACAGGTACTGATAATTATGATGTTGGCGGTAAAACTTATACCGTGAATG  
 TGGAGAGCGCGCGGTTAAGAATGATGCTAATAAAGATGTTTTTGAAGCGCAGCTGATG  
 GATCGCTGACGACAGTAGTGATACTAAAGTATCCGGTGAAAGTATTGATGCAACAGAAC  
 TAGCGAAACTTGCAATAAAATTAGCTGACAAAGGCTCCATTGAATACAAAGGCGATTACAT  
 TTAATAACAACACTGGCGCAGAGCTTGATGCTAATGGTAAAGGTGTTTTGACCGCAAATA  
 TTGATGGTCAAGATGTTCAATTTACTATTGACAGTAATGCACCCACGGGTGCCGCGCAAA  
 CAATAACTACAGACACAGCTGTTTTACAAAAACAGTGCGGGCCAGTTTCACTACAAAAAG  
 TGGAAAAATAAAGCCGCAACACTCTCTGATCTGGATCTTAATGCAGCCAAAGAAAACAGGTA  
 GCACCTTTAGTTGTAATGGCGCCACCTACAATGTGACGCGCAGATGGTAAACGGTAACGTG  
 ATACTACTCTCTGGTGCCCTAAAGTGATGTATCTGAGCAAAATCAGAAGGTGGTAGCCGGA  
 TTCTGGTAAACGAAGATGCAGCAAAATCGTTGCAATCTACCAACCAACCCGCTCGAAACTA  
 TCGACAAGGCATTGGCTAAAGTTGACAATCTGCGTTCTTGACCTCGGTGACGTACAAAAACC  
 GTTTCGACTCTGCCATCACCACCTTGCGCAACCCGTAACAAACCTGTCTTCTGCCCCTGA  
 GCGGTATCGAAGATGCTGACTACGCGACCGAAGTGCTAACAATGCTCTGTCGCGCAGATCC  
 TGCAACAAGCGGTACCTCTGTTCTGGCGCAG

Figure 14

37/96

ATGGCACAAAGTCATTAATACCAACAGCCTCTCG  
 CTGATCACTCAAAATAATATCAACAGAACCAGTCTCGCTGTGAGTTCTATCGAGCGT  
 CTGTCTTCTGGCTTGGCTATTAAACAGCGCGAAGGATGACGCCGCGGTCAGGCGGATTGCT  
 AACCGTTTTACTTCTAAACATTAAAGGCTGACTCAGGCTGCACGTAAACGCCAACGACGGT  
 ATTTCCGTTGCACAGACCACTGAAGCGCGCTGTCCGAAATTAACAACAACCTACAGCGT  
 ATTCGTGAACCTGACGGTTCAGGCTTCTACCGGACTAATCCGATTCCGATCTGGACTCC  
 ATTCAGGACGAAATCAAATCCCGCTGTGGACGAAATTGACCGGTATCCGGCCAGACCCAG  
 TTCAACGGCGTGAACGTGTCTGTCCAAAGATGGCTCGATGAAATTCAGGTCCGCGCGAAC  
 GATGGCGAAACGATTACTATTGATCTGAAGAAAAATGACTCTGATACGCTGAATCTGGCT  
 GGTTTTAACGTTAACCGTAAAGGTTCTGTAGCGAATACAGCTGCGACAAGCGACGATTTA  
 AAACCTGGCTGGTTTCACTAAGGGCACACAGATACCAATGGCGTGACCGGTATACAAAC  
 ACAATTAGTAATGACAAAGCCAAAGCTTCCGATCTGTTAGCTAATATCACCGATGGATCA  
 GTGATCACTGGGGAGGGGCAAACGCTTTTGGCGTGGCTGCAAGAATGGTTACACCTAT  
 GATGCGACGAAGTAAATCTTATAGTTTGTGTCAGATGGTGCCGATTCAAGCGAAGACGTTA  
 AGCATCATTAATCCAAACACCGGTGATTCTGTCGACGGCGACAGTGACTATTGGTGGTAAA  
 GAGCAGAAAGTTAATATTTCCAGGATGGAAAAATTAAGTCCGCGAGATGATAATGCGAGC  
 CTGTATTTAGATAAACAGGGAACCTTGACAAAAACGAAATACAGGTAACGATACCGCAGCG  
 ACTTGGGATGGTTTAATTTCCACAGCGATTCTACCGGTGCGGTTCCAGTTGGGGTTGCA  
 ACTACAATTACAATTACTTCTGGTACAGCTTCCGGAATGTCTGTTCAAGTCCGCGAGGCA  
 GGAATTCAGACCTCAACAAATTCTCAGATTCTTGAGGTGGTGCAATTGCGGCTAAGGTA  
 AGTATTGAGGGAGGCGCTGCTACAGACATTTTGGTAGCAAGTAATGGAAACATAACAGCG  
 GCTGATGGTAGTGCACTTTATCTTGATGCGACTACTGTGGATTCACTACAACGGCTGGA  
 GGAAATACAGCTGCTTCGTAGATAATTTAATTGCTAACAGTAAGGATGCTACCTTAACC  
 GTAACCTCAGGTACCGGCCAGAACACTGTTTATAGCACAACAGGAAGTGGCGCTCAGTTC  
 ACCAGTTTAGCAAAAGTAGACACAGTCAATGTCAACCAACGACATGTCAAGTCCGCAAGGT  
 ATGGCAAATCTGACAAAAAGCAATTTTACCATTGATATGGGCGGTACAGGTACAGTAAC  
 TACACAGTTTCCAATGGGGATGTGAAAGCTGCTGCAAAATGCTGATGTTTATGTGCAAGAT  
 GGTGCACCTTCAGCCAATGCTACAAAGATGTAACTACTTTGAACAAAAAATGGGGCT  
 ATTACCAACAGCACCGGTGGTACCATCTATGAACAGCTGATGGTAAGTTAACACAGAA  
 GCTACTACTGCATCCAGTTCCACCGCCGATCCCTGAAAGCTCTGGACGAAGCCATCAGC  
 TCCATCGACAAATCCGCTCCCTCCGTCGGTGCGTAACACCGTCTGGAATTCGCGGGT  
 ACCAACCTGAACAAACCACTACCAACCTGTCCGAAGCGCAGTCCCGTATTCAAGACGCC  
 GACTATGCGACCGAAGTGTCCAACATGTGCAAGCGCAGATCATCCAGCAGGCCGGTAAC  
 TCCGTGCTGCGCAAAGCTAACAGGTAACCGCAGCAGGTTCTGTCTCTGTCGAGGTTAA

Figure 15

38/96

ATGGCACAAAGTCATTAATACCAACAGC  
 CTCTCGCTGATCACTCAAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATC  
 GAGCGTCTGTCTTCTGGCTTGCGTATTAACAGCGCGAAGGATGACGCCGCGGTCAGGCG  
 ATTGCTAACCGTTTTACTTCTAACATTAAGGCGCTGACTCAGGCTGCACGTAACGCCAAC  
 GACGCTATTCTGTTGCGCAGACCCGAAGGCGCGCTGTCCGAAATTAACAACAATTA  
 CAGCGTGTGCGTGAGCTGACTGTTTCTGAGCGACCCAGGCTACTAACTCTGAGTCTGACCTG  
 TCTTCTATCCAGGACGAAATCAAATCTCGCCTGGAAGAGATTGATCGTGTTCAGTCTG  
 ACTCAATTTAAGCGCGTGAAATGTTTGGCTAAAGATGGGAAATGAACATTGAGTTGGG  
 GCAAATGATGGACAGACTATCACTATTGATCTGAAAAGATCGATTCACTCACTAAAC  
 CTCTCCAGTTTTGATGCTACAACTTGGGCACCACTGTTAAAGATGGGGCCACCATCAAT  
 AAGCAAGTGGCAGTAGGTGCTGGCGACTTTAAAGATAAAGCTTCAGGATCGTTAGGTACC  
 CTAAATTAGTTGAGAAAGACGGTAAGTACTATGTAATGACACTAAAAGTAGTAAGTAC  
 TACGATGCCGAAGTAGATACTAGTAAGGGTAAATTAACCTCAACTCTACAAATGAAAGT  
 GGAAGTACTCCTACTGCAGCGCAGGAAGTAAGTAACTACTGTTGGCCGCGATGTAATTTGGAT  
 GCTTCTGCACTTAAAGCCAAACCAATCGCTTGTCTGTTATAAAGATAAAGCGGCAATGAT  
 GCTTATATCATTCAGACCAAGATGTAACAACATAATCAATCAACTTCAATGCGCGTAAT  
 ATCAGTGATGCTGGTGTCTTATCTATTGGTGCATCTACAACCGCGCCAAGCAATTTAACA  
 GCTAACCCGCTTAAGGCTCTTGATGATGCAATGCACTCTGTTGATAAATTCGCGTCTTCT  
 TCGGTGCGCTTCAGAACCGTCTGGATTCTGCCATTGCCAACCTGAACAACACCACTACC  
 AACCTGTCTGAAGCGCAGTCCCGTATTCAGGACGCTGACTATGCGACCGAAGTGTCCAAC  
 ATGTGCAAAAGCGCAGATTATCCAGCAGGCGCGTAACCTCGGTGCTGGCAAAAGCCAACCAAG  
 GTACCGCAGCAGGTTCTGTCTCTGCTGCGAGGGTTAA

Figure 16

39/96

AACAAATCTCAGTCTTCTCTGAGCTCCGCCAT  
 TGAACGTCTCTCTCTGGCCTGCGTATTAAACAGTGTAAAGATGACGACGAGGTGAGGC  
 GATTGCTAACCGTTTTACAGCAAATATTAAAGGTCTGACTCAGGCTTCCCGTAACGCGAA  
 TGATGGTATTCTGTTGCGCAGACCACTGAAGGTGCGTGAATGAAATTAACAAACCT  
 GCAGCGTGACGTAACTGACTGTTTCAGGCACTAACGGTACTAACTCTGACAGCGATCT  
 TTCTTCTATCCAGGCTGAAATTACTCAACGTCTGGAAGAAATTGACCGTGTATCTGAGCA  
 AACTCAGTTTAAAGGCGTGAAGTCTTGTCTGAAATAATGAAATGAAATTCAGGTTGG  
 TGCTAATGATGGTGAACCATCACTATCAATCTGGCAAAAATGATGCGAAAACTCTCGG  
 CCTGGACGGTTTTAATATCGATGGCGCGCAGAAAGCAACTGGCAGTGACCTGATTCTAA  
 ATTTAAAGCGACAGGTACTGATAACTATGATGTTGGCGGTGATGCTTATACTGTTAACGT  
 AGATAGCGGAGCTTTAAAGATACTACAGGGAATGATATTTTGTGTAGTCAGCAGATGG  
 TTCCTGACAACTAAATCTGACACAAACATAGCTGGTACAGGATTGATGCTACAGCACT  
 CGCAGCAGCGGCTAAGAATAAAGCACAGAATGATAAATCACGTTTAAATGGAGTTGAATT  
 CACRACAACTGCAGCGGATGGCAATGGGAATGGTGTATATTCTGCAGAAATGATGG  
 TAAAGTCAGTGACATTTACTGTGACAGATGCTGACAAAAAGCTTCTTTGATTACAGTGA  
 GACAGTTTACAAAAATAGCGCTGGCCTTTATACGACAACCAAAGTTGATAACAAGGCTGC  
 CACACTTCCGATCTTGATCTCAATGCAGCTAAGAAAAACGGAAGCACGTTAGTTGTTAA  
 CGGTGCAACTTACGATGTTAGTGACAGATGGTAAACGATAACGGAGACTGCTTCTGGTTAA  
 CAATAAGTCATGTATCTGAGCAAAATCAGAAGGTGGTAGCCGATTCTGGTAAACGAAGA  
 TGCAGCAAAATCGTTGCAATCTACCACCAACCGCTCGAAACTATCGACAAAGCATTTGGC  
 TAAAGTTGACAATCTGCGTCTTGACCTCGGTGCAGTACAAAACCGTTTCGACTCTGCTAT  
 CACCAACTTTGGCAACCCGTAACCAACCTGTCTTCTGCGCGTAGCCGTATCGAAGATGC  
 TGACTACGCGACCGAAGTGCTAACATGTCTCGTGCGCAGATCCTGCAACAAGCGGCTAC  
 CTCTGTTCTGGCGCAG

Figure 17

40/96

ATGGCACAAAGTCATTAAATACCAACAGCCTCTCGCTGATCA  
 CTCAAAAATATATCAACAAGAACCAGTCTGCGCTGTGCGAGTTCATCGAGCGTCTGTCTT  
 CTGGCTTGCCTATTAAACAGCGCGAAGGATGACGAGCGGGTCAGGCGATTGCTAACCGTT  
 TCACCTCTAAACATTAAGGCTGACTCAGGCGGCCCGTAACGCCAACGACGGTATCTCCG  
 TTGGCGCAGACCAACGAAAGCGCGCTGTCCGAAATCAACAACAACCTTACAGCGTATCCGCG  
 AACTGACGGTTCAGGCTTCTACCGGACTAATCCGATTGCGATCTGGACTCCATTACAGG  
 ACGAAATCAAAATCCCGTCTGGACGAAATTGACCGCGTATCTGCGCAGACCCAGTTCAACG  
 GCGTGAACGCTACTGCGCAAGACGGTTCAATGAAAATTCAGGTTGGTGCGAATGACGCGCC  
 AGACTATCAGGATGATCTGAAGAAAATTGACTCAGATACGCTGGGGCTGAATGGTTTTA  
 ACGTGAATGGTTCGGGTACGATAGCCAATAAAGCGGCGACCATTAGCGACCTGACAGCAG  
 CGAAAAATGGATGCTGCAACTAATACTATAACTACAACAAATAATGCGCTGACTGCATCAA  
 AGGCGCTTGATCAACTGAAGAATGGTGACACTGTTACTATCAAAGCAGATGCTGCTCAAA  
 CTGCCAGCGTTTTATACATACAATGCATCAGCTGGTAACCTTCTCATTCAAGTAATGTATCGA  
 ATAATACTTTCAGCAAAAGCAGGTGATGTAGCAGCTAGCCTTCTCCCGCCGGCTGGGCAAA  
 CTGCTAGTGGTGTATTAAGCAGCAGCGGTGAAGTGAACCTTTGATGTTGATGCGAATG  
 GTAAAATCACAATCGGAGGACGAAAGCATATTTAACTAGTGTGGTAACTTAACTACAA  
 ACGATGCTGGTGGTGGACTGCGCTACGCTTGATGGTTTTATTCAAGAAAGCTGGTGATG  
 GTCAATCAATCGGTTTTAAGAAGACTGCATCAGTCAGTATGGGGGGAACAACTTATAACT  
 TTAACACGGGTGCTGATGCTGATGCTGCACTGCTAACGCGAGGGTATCGTTCACTGATA  
 CAGCTAGCAAAAGAAACCGTTTTAAATAAAGTGGCTACAGCTAAACAAGGCAAGCAGTTG  
 CAGCTGACGGTGATACATCCGCAACAATTACCTATAAATCTGGCGTTCAGACGTATCAGG  
 CTGTATTTGCGCGCAGGTGACGGTACTGCTAGCGCAAAATATGCCGATAAAGCTGACGTTT  
 CTAATGCAACAGCAACATACACTGATGCTGATGGTGAATGACTACAATTGGTTCATACA  
 CCACGAAGTATTCAATCGATGCTAACAAACGGCAAGGTAACCTGTTGATTCTGGAACGTGTA  
 CGGGTAAATATGCGCGAAAGTAGGGGCTGAAGTATATGTTAGTGCTAATGCTACTTTAA  
 CAACAGATGCAACTAGCGAAGGCACAGTAACAAGATCCACTGAAAGCTCTGGATGAAG  
 CTATCAGCTCCATCGACAAATTCGTTCTCCCTGGGTGCTATCCAGAACCGTCTGGATT  
 CCGCAGTCAACCACTGAACAACCACTACCAACCTGTCCGAAGCGCAGTCCCGTATTC  
 AGGACGCCGACTATGCGACCGAAGTGTCACACATGTGAAAGCGCAGATCATTACAGCAG  
 CCGGTAACTCCGTGCTGGCAAAAGCCAACAGGTACCGCAGCAGGTTCTGCTCTGCTGCG AGGGTTAA

Figure 18

41/96

ATGGCACAAAGTCATTAAATACCAACAGCCTCTCGTGATCACTAAAAATA  
 ATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATCGAGCGTCTGTCTTCTGCGTTGC  
 GTATTAAACAGCGCGAAGGATGACGCGCGCAGGTGAGCGGATTGCTAACCGTTTTACTTCTA  
 ACATTAAAGGCGCTGACTCAGGCTGCACGTAACGCCAACGACGGTATTTCCGTTGCGCAGA  
 CCACTGAAGGTGCGCTGTCCGAAATCAACAACACTTACAGCGTATTCTGTAGCTGACGG  
 TTCAGGCTTCTACCGGGACTAACTCCGATTCTGACCTGGACTCCATCCAGGACGAAATCA  
 AGTCTCGTCTGGACGAAATTGACCGCGTATCCGGTCAGACCCAGTTCAACGGCGTGAACG  
 TGCTGGCGAAAGACGGTTTCGATGAAAATTACGGTTGGTGCGAATGACGGCCAGACTATCA  
 CGATTGATCTGAAGAAATTGACTCAGATACGCTGGGGCTGAGTGGGTTTAATGTGAATG  
 GTGGCGGGGCTGTTGCTAACTGCTGCATCTAAAGCTGACTTGGTAGCTGCTAATGCAG  
 CTGTGGTAGGCAACAAATATACTGTGAGTGCGGGTTACGATGCTGCTAAAGCGTCTGATT  
 TGCTGGCTGGAGTTAGTGATGGTGATACTGTTCAAGCAACCATTAATAACGGCTTCGGAA  
 CGGCGGCTAGTGCAACGAATTACAAGTATGACAGTGCAAGTAAAGTCTTACTCTTTTGATA  
 CCACAACGGCTTCAGCTGCCGATGTTTCAGAAATATTTGACCCCGGGCTGGTGATACCG  
 CTAAGGGCACTATTACTATCGATGGTTCTGCACAGGATGTTTCAGATCAGCAGTGATGGTA  
 AAATTACGTCAAGCAATGGAGATAAACTTTACATTGATACAACTGGGCGCTTAACGAAAA  
 ACGGCTTTAGTGCTTCTTTGACTGAGGCTAGTCTGTCCACACTTCAGCGCAATAATAACCA  
 AAGCGACAACCATTGACATTGGCGGTACCTCTATCTCCTTTACCGGTAATAGTACTACGC  
 CGAACACTATTACTTATTCAGTAACAGGTGCAAAAGTTGATCAGGCAGCTTTTCGATAAAG  
 CTGTATCAACCTCTGGAAACGATGTTGATTTCACTACCGCAGGTATAGCGTCGACGGCG  
 CAACTGGCGCTGTAACAAAAGGTGTTGCTCCGGTTTATATTGATAACAACGGGCGTTGA  
 CCACATCTGATCTAGATTTTATCTACAGGATGATGGTTGAGTGACTAACGGCAGCG  
 GTAAGGCGAGTTTATAAAGATGCTGACGGTAAATTGACGACAGATGCTGAAACTTAAAGCTG  
 CAACCAACCGCGATCCCCGAAAAGCTGTGACGAAGCCATCAGCTCCATCGACAAAATTC  
 GCTCCTCCCTCGGTGCGGTGCAGAACCGTCTGGATTCCGCGGTACCAACCTGAAACAACA  
 CCATACCAACCTGTCTGAAGCGCAGTCCCGTATTGAGGACGCTGACTATGCGACCGAAG  
 TATCCAACATGTCGAAAGCGCAGATCATCCAGCAGGCGGTAACCTCCGTGCTGGCAAAAG  
 CTAACCAAGTACCACAGCAGGTTCTGTCTCTGCTGACAGGTTAA

Figure 19



42/96

ATGGCACAAGTCATTAAATACCAACAGC  
 CTCTCGTGATCACTCAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATC  
 GAGCGTCTGTCTCTGGCTTGGCTATTAAACAGCGCGAAGGATGACCCGCGAGGTGAGGCG  
 ATTGCTAACCCTTTTACTTCTAACATTAAAGGCGCTGACTCAGGCTGCACGTAAACGCCAAC  
 GACGGTATTTCTGTTGCACAGACCACTGAAGGCGCGCTGTCGAAATCAACAACAACCTTA  
 CAGCGTGTGCGTGAACAGCGTTCAAGGCAACCAACCGGTACCAACTCCAGTCTGACCTG  
 GACTCTATCCAGGACGAAATTAATCCCGTCTGGACGAAATGTATCGGTATCCGGTCAG  
 ACCCAGTTCAACGGCGTGAACGTGCTGGCAAAGACGGTTCATGAAAATTCAGGTTGGC  
 GCGAACGATGGCCAGACCATCACTATCGACCTGAAGAAGATTGACTCTTCTACCTTGAAC  
 CTGACAGGTTTTAACGTTAACGGTTCTGGTTCTGTTGGCGAATACTGCAGCAACTAAAGCT  
 GATTTAACCGCTGCTCAACTCTCTGCACCGGGTGCAGCAGACGCAATGGTACAGTTACT  
 TATACTGTCAGTGTGTTTATAAAGAAATCCACTGCTGCAGATGTTATTGCTAGCATCAA  
 GACGGCAGTGTCCGACTTCTGCAATTACTGCAACCATTAATAATGGCTTCGGTGATTCC  
 AGTGGCGTGACTTCCAATGACTATACTTATGACCAGCAAAAGGCGACTTCACTTACGAC  
 GTAGCTTCAAGCGCCAAATAACTGCTGCCAGGTTGAGTCTTCTGACGCCGAAAGCA  
 GGTGATACCGCAAAATCTGAAAGTAACCGTTGGTACGACATCGGTTGATGTCGTTCTGGCC  
 AGTGATGGTAAGATTACAGCAAAAGATGGTTCTGCATTATATCGACAGTACAGGTAAC  
 CTGACTCAGAACAGTGTCTGGCTTGACCTCTGCTAAACTGGCTACTCTGACTGGCCTTCAG  
 GGCTCTGGTGTGCTTCAACCATCACTACTGAAGATGGCACTAATATTGATATTGCTGCT  
 AACGGTAATATTGCTCTGACCGGTGTTGCTATCAGTGTGATTCTCTGCAGTCAGCGACT  
 AAATCTACGGGCTTTACTGTTGGTACTGGCGCTACAGGTCTGACCGTAGGTACTGATGGT  
 AAAGTGACTATCGCGGGACTACTGCTCAGTCTTACACCAGCAAGATGGTTCCTGACT  
 ACTGATAACACCACTAACTGTATCTGCAGAAAGATGGCTCTGTAAACCAACGGTTCAAGT  
 AAAGCGGTCTATGTAGAAGCGGATGGTGATTTCACTACCGACGCTGCAACCAAGCGCGCA  
 ACCACCAACCGATCCGCTGAAAGCCCTGGATGAGGCAATCAGCCAGATCGATAAGTTCCTG  
 TCATCCCTGGGTGCTATCCAGAACCGTCTGGATTCCGCGGTCAACCACTGAACAACACC  
 ACTACCAACCTGTCTGAAGCGCAGTCCCGTATTACAGGACGCCGACTATGCGACCGAAGTG  
 TCCAACTGTGAAAGCGCAGATCATTCAGCAGGCGGTAACTCCGTGCTGGCAAAGCC  
 AACCAGGTACCGCAACAGGTTCTGCTCTGCTGCAGGGCTAA

Figure 20

43/96

ATGGCACAAGTCATTAATACCAACAGCCTCTCGCTGATCAC  
 TCAAAATAATATCAACAAGAACCAGTCTGCGCTGTGAGTTCTATCGAGCGTCTGTCTTC  
 TGGCTTTCGTATTAAACAGCGCGAAGGATGACGCCGAGGTGAGGCGATTGCTAACCGTTT  
 TACTTCTAAACATTAAAGGCCTGACTCAGGCTGCACGTAACGCCAACGACGGTATTTCGTG  
 TGCACAGACCACTGAAGGCGCGCTGTCCGAAATCAACAACACTTACAGCGTGTGGGTGA  
 ACTGACCGTTTCAAGCAACCACCGGTACCAACTCCCGAGTCTGACCTGGACTCTATCCAGGA  
 CGAAATTAAATCCCGTCTGGACGAAATTGATCGCGTATCCGGTCAAGCCAGTTTCAACGG  
 CGTGAAACGTGCTGGCAAAAGACGGTTCATGAAAATTACGGTTGGCGCGAAGCATGGCCA  
 GACCATCACTATCGACCTGAAGAAGATTGACTCTTCTACCTTGAACCTGACAGGTTTAA  
 CGTTAACGGTTCTGGTTCCTGTGGCGAATACTGCAGCAACTAAAGCTGATTTAACCGCTGC  
 TCAACTCTCTGCACCCGGGTGCAGCAGACGCAAAATGGTACAGTTACTTATACTGTCACTGC  
 TGGTTATAAAGAAATCCACTGCTGCAGATGTTATTGTAGCATCAAAGACGGCAGTGTCTCC  
 GACTTCTGCAATTACTGCAACCATTAAATAATGGCTTCGGTGATCCAGTGCGCTGACTTC  
 CAATGACTATACTTATGACCCAGCAAAAGGCGACTTCACTTACGACGTAGCTTCAAGCGC  
 CAATAACTGCTGCGCCAGGTTTCACTCCTTCTGACGCCGAAAGCAGGTGATACCGCAA  
 TCTGAAAGTAACCGTTGGTACGACATCGGTTGATGTCGTTCTGGCCAGTGATGGTAAGAT  
 TACAGCAAAAGATGGTTCTGCATTATATATCGACAGTACAGGTAACTGACTCAGAACAG  
 TGCTGGCTTGACCTCTGCTAAACTGGCTACTCTGACTGGCCCTTCAAGGCTCTGGTGTTC  
 TTCAACCATCACTACTGAAGATGGCACTAATATTGATATTGCTGCTAACCGGTAATATTGG  
 TCTGACCCGGTGTTCGTATCAGTGTGATTCTCTGCAGTCAGCGACTAAATCTACGGGCTT  
 TACTGTTGGTACTGGCGCTACAGGTCTGACCGTAGGTACTGATGGTAAAGTGACTATCGG  
 CGGGACTACTGCTCAGTCTTACACCCAGCAAAAGATGGTTCCCTGACTACTGATAACACCAC  
 TAAACTGTATCTGCAGAAAGATGGCTCTGTAACCAACCGGTTCAAGGTAAGCGGCTATGT  
 AGAAGCGGATGGTGATTTCACTACCGACGCTGCAACCAAGCCGCAACACCACCGATCC  
 GCTGAAAGCCCTGGATGAGGCAATCAGCCAGATCGATAAGTTCCGTTTCATCCCTGGGTGC  
 TATCCAGAACCGTCTGGATTCCCGGTCACCAACCTGAACAACCACTACCAACCTGTG  
 TGAAGCGCAGTCCCGTATTTCAGGACGCCGACTATGCGACCGAAGTGTCCAACATGTGAA  
 AGCGCAGATCAATCAGCAGGCCGGTAACTCCGTGCTGGCAAAAGCCAAACAGGTACCGCA  
 ACAGGTTCTGTCTCTGCTGCAGGGCTAA

Figure 21

44/96

GCGCTGTCGACTTCTATCGAGCGCCTCTCTTCTGGTCTGCGTATTACAGCGCTAAA  
 GATGACGCTCGGGCCAGGCGATTGCTAACCGCTTCACTTCTAACATCAAAGTCTGACT  
 CAGGCCGCACGTAACGCCAACGACGGTATTTCTCTGGGCGCAGACGGCTGAAGGCGCGCTG  
 TCAGAGATTAAACAACAATTGACGCGTATTCGTGAACGTGACCGTTACAGGCCCTTACCGGC  
 ACGAACTCTGATTCCGACCTGTCTTCTATTGAGACGAAATCAAATCCCGTCTTGATGAA  
 ATTGACCGTGTATCTGGTCAGACCCAGTTCAACGGTGTGAACGTGCTGTCGAAAAACGAT  
 TCGATGAAGATTGAGATTGGTGCCAATGATAACCGACGATCAGCATTGGCTTGCAACAA  
 ATCGACAGTACCACTTTGAATCTGAAAGGATTTACCGTGTCCGGCATGGCGGATTTTCAGC  
 GCGGCGAACTGACGGCTGCTGATGGTACAGCAATTGCTGCTGCGGATGTCAAGGATGCT  
 GGGGTAAACAAGTCAATTTACTGTCTTACACTGACACCGGTCTAACAGTACTAAATAT  
 GCGGTGCTGATTCTGCAACCGGTAATACATGGCAGCCACTGTAGTCATTACAGTACG  
 GCGGCGGCGTAACTGTTGGTGCAACGGAAGTGGCGGGAGCCGCTACAGCCGAACCGTTA  
 AAAGCACTGGATGCGCGCAATCGCTAAAGTCGACAAATCCCGTCTCCCTCGGTGCCGTT  
 CAAAACCGTCTGGATTCTGCGGTCAACCACTGAACAACACCCACCAACCTGTCTGAA  
 GCGCAGTCCCGTATTGAGACGCCGACTATGCGACCGAAGTGTCCAACATGTCGAAAGCG  
 CAGATTATCCAGCAGGCG

Figure 22

45/96

ATGGCACAGTCATTAATACCAACAGCCTCTCGCTGATCACTCAAAATA  
 ATATCAACAAGAACCCAGTCTGCGCTGTGCGAGTCTATCGAGCGCTGTCTCTTCTGGCTTGC  
 GTATTAACAGCCGCAAGGATGACGCCGCGAGGTCAAGCGATTGCTAACCGTTTACTTCTA  
 ATATTTAAAGCCCTGACTCAGGCTGCACGTAAACGCCAATGACGGTATTTCTGTTGACACAGA  
 CCACTGAAGGCGCGCTGTCCGAAATCAACAACTTACAGCGTATTTCGTGAACCTGACGG  
 TTCAGGCCACTACAGGGACTAACTCCGATTCTGACCTGGACTCCATCCAGGACGAATCA  
 AATCTCGTCTGGACGAAATTGACCGGTATCCGGTCAGACCCAGTTCAACGGCGTGAACG  
 TGCTGTCCAARGATGGTTCAATGAAAATTCAAGTCCGCGCAAAATGATGGTGAACCATCA  
 CGATTGATCTGAAGAAAATTGACTCTGATACGTGAATCTGGCTGGTTTTAAACGTGAATG  
 GCGAAGGTGAAACAGCCAATACTGCTGCAACACTTAAAGATATGGTTGGTTTAAAACTCG  
 ATAATACGGGGGTCACCTACAGCTGGAGTTAATAGATATATTGCTGACAAAGCCGTCGCAA  
 GTAGCACGGATATTTGAATGCGGTAGCTGGTGTGATGGCAGTAAAGTTCCACGGAGG  
 CAGATGTTGGTTTTGGTGCAGCTGCCCTGGTACGCCAGTGGAAATATACTTATCATAAAG  
 ATACTAACACATATACGGCTTCTGCTTCAGTTGATGCGACTCAACTGGCGGCATTCCTGA  
 ATCCTGAAGCGGTTGGTACCACTGCTGCAACAGTAAGTATTGGCAACGGTACAACAGCTC  
 AAGAGCAAAAAGTCATTATTGCTAAAGATGGTTCTTTAACTGCTGCTGATGACGGTGCCG  
 CTCTCTATCTTGATGATACTGGTAACTTAAGTAAAACTAACGCAGGCACTGATACTCAAG  
 CTAACCTGTCTGACTTAATGGCAACCAATGCTAATGCCAAAACAGTCATTACAACAGATA  
 AAGGTACATTTACTGCTAATACGACAAAAGTTTGATGGGGTAGATATTTCTGTTGATGCTT  
 CAACGTTTGCTAACGCCGTTAAAAATGAGACTTACACTGCACTGTTGGTGAACTTTAC  
 CTGCGACATATACAGTCAATAATGGCACTGCTGCATCAGCGTATTAGTCGATGGAAG  
 TGAGCAAACTCCTGCCGAGTATTTTGCTCAAGCTGATGGCACTATTACTAGTGGTGAAG  
 ATGCCGCTACCAGTAAAGCTATCTATGTAAGTGCCAATGGTAACTTAACGACTAATACAA  
 CTAGTGAATCTGAAGCTACTACCAACCCGCTGGCAGCATTTGGATGACGCTATCCGCTCTA  
 TGCACAAATTCGGTTCTTCCCTGGGTGCTATCCAGAACCGTCTGGATTCCGCGAGTCAACA  
 ACGTGAACAACACCACTACCAACCTGTCTGAAGCGCAGTCCCGTATTACAGGACGCCGACT  
 ATGCGACCGAAGTGTCCAACATGTGAAAGCGCAGATCATTACAGCAGGCCGTTAACTCCG  
 TGCTGGCAAAAGCAACCAAGTACCGCAGCAGGTTCTGTCTCTGCTGCAGGGTTAA

Figure 23

46/96

ATGGCACAGTCATTAATACCAACAGCCTCTCGCTGATCACTCAAAATATAT  
 CAACAAGAACCAAGTCTGCGCTGTGAGTTCTATCGAGCGTCTGTCTTCTGGCTTGGTAT  
 TAACAGCGCGAAGGATGACGCGCAGGTCAGGCGATTGCTAACCGTTTTACTTCTAACAT  
 TAAAGGCTGACTCAGGCTGACGTAACGCCAACGACGGTATTCTGTTGCGCAGACCA  
 TGAAGGCGCGCTGTCCGAAATTAACAACAATTACAGCGTATTCTGTAAGTACGCGTTCA  
 GCGCAGCAGCGGAATTAACCTCCACTCTGACCTGGACTCCATCCAGGACGAAATCAATC  
 CGGTCTTGACGAAATTGACCGCGTATCTGGTCAGACCCAGTTCAACGGCGTGAACGTGCT  
 GTCTAAAGATGGCTCGATGAAATTCAGGTGCGCGCAACGATGGCGAAACGATTACTAT  
 TGATCTGAAGAAATTGACTCTGATACGCTGAATCTGGCTGGTTTAAACGTTAACGTTAA  
 AGGTTCTGTAGCGAATACCGCTGCGACTACAGATAATCTGACATTGGCTGGTTTTACAGC  
 GGGTACTAAAGCTGCTGATGGCACCGTAACTTATAGCAAAATGTCCAGTTTGCCGCCGC  
 GACTGCAAGCAATGTACTGGCTGCTGCTAAAGATGGCGACGAAATTACGTTCTGCTGGTAA  
 TAACGGCACAGGTATAGCTGCAACTGGGGGACTTATACCTTATCATAAAGACTCTAACTC  
 ATACAGCTTTAGCGCAACGGCTGCATCTAAAGATTCTCTGTTGAGCACACTGGCACCAAA  
 CGCTGGCGTACATTTACCGCTAAAGTGAATTTGGTTCTAAATCGCAAGAAAGTTAAAGT  
 TAGCAAGATGGTACGATTACATCCAGCGATGGTAGGCGCTGTATTTAGATGAGAAGGG  
 CAACCTGACCCAAACAGGTAGTGGCAACCAAGCTGCAACCTGGGATAACCTGATGGC  
 CAATACAGATACTACAGGCAAGATGCTTATGTTAACTCTGCGGCAGCAGCTGTTGGGAC  
 AGTAATCGAAGCAAAAGGAATGACCATCACTTCTGCTGGTGGTAAATGCTCAGGTGTTAA  
 AGACGCGGCTTATAATGCGCATATGCGACCTCAATTACTCTGGTACTCCGGGTGATGC  
 GGGAGCGCGGGAGCCGCTGCAACTGCGGGTAATGCGCGGTGGGAGCGCTGGGCGCAAC  
 GGCAGTTGATAATACCGCGCAGATGTTGCGGATATCTCTATCTCAGCTTCGCAATGGC  
 GAGCATCTCTCAGGATAAAGATTTCACCTTAAGTGATGGTAGTGATACTTACAACGCTGAC  
 CAGCAATGCTGTCACTATCAATGGCAAGCAGCAACATTGATGACAGCGGCGCAATCAC  
 AGACCAAAACAGTAAAGTTGTCAATTATTTGCTCATACTAACGGTAGCGTACTAACGA  
 TACAGGCTCCACTATTTATGCGACAGAAGATGGTAGCCTGACCAACGATGACGCAACCA  
 AGCCGAAACACCGCGATCCCTGAAAGCTCTGGACGAAGCCATCAGTCCATCGACAA  
 ATTCCGCTCTCCCTCGGTGCGGTGCAAAACCGTCTGGATTCCGCGGTCAACACCTGAA  
 CAACACCACCACCACTGTCTGAAGCGCAGTCCCGTATTGAGGACGCCGACTATGCGAC  
 CGAAGTGTCAACATGTCGAAAGCGCAGATTATCCAGCAGCGCGGTAACTCCGTGCTGGC  
 AAAAGCTAACAGGTACCAACAGCAGGTCTGTCTCTGCTGACGGTTAA

Figure 24

47/96

ATGGCACAAAGTCATTAATACCAACAGCCTCTCGCTG  
 ATCACTCAAAATAATATCAACAAGAACAGCTCTGCGCTGTCGAGTTCTATCGAGCGTCTG  
 TCTTCTGCGCTTGGGTATTAAACAGCGCGAAGGATGACGCCGAGTTCAGGCGATTGCTAAC  
 CGTTTTACTTCTAACATTAAAGCGCTGACTCAGGCGGCCGCTAACGCCAACGAGCGTATT  
 TCTGTTGCGCAGACCCGGAAGGCGCGCTGTCCGAAATTAACAACAACCTTACAGCGTGTG  
 CGTGAGCTGACTGTTTCAAGCGACACCCGCTACCACTCCCAGTCTGATCTGGACTCTATC  
 CAGGACGAAATCAAAATCCCGTCTGGACGAAATTGACCGCGTATCCGGTCAGACCCGAGTTC  
 AACGCGGTGAACGTGCTGGCAAAAGACGTTCCATGAAAATTCAGGTGGCGCGAATGAT  
 GGCCAGACCATCACTATCGACCTGAAGAAGATTGACTCTTCTACGTTGAAACTGACTGGT  
 TTTAACGTGAATGGTTCTGGTCTGTGGCGAATACTGCGCGCACTAAAGCGGATTGGCT  
 GCTGCTGCAATTGGTACCCCTGGGCGACGAGATTCTACAGGTGCCATTGCTTACACAGTA  
 AGTGCTGGGCTGACTAAACTACAGCCGAGATGACTGTCTAGCCTCGCTGATGGTAGC  
 ACTAATTACGCCACAGGCGTGAAAAATGGCTTTGCTGCAGGAGCCACTTCCAATGCCAT  
 AAAGTTAACAAAGATAATAATACATTTACTTATGACACGACTGCTACGACAGCTGAGCTG  
 CAGTCTTACCTGACTCCGAAAGCGGCGCACTGCAACATTCAGTGTGAAATTTGGTGGT  
 ACTACACAAGACGTGCTGTCTGTCAGTGATGGCAAACTCACTGCTAAGGATGGCTCTAAG  
 CTTTACATTGATACAACCTGGTAATTTAACTCAGAATGGTGGTAATAACGGTGTGGAAACA  
 CTCGCGGAAGCGACTCTGAGTGGTTTGTCTGTGAACAAAATGGTTTAAACGGCTGTATAA  
 TCCACAATTACTACAGCTGATAACACTTTCGATTGTACTGAATGGTTCAAGCGATGGTACT  
 GGTAAATGCTGGTACTGAAGGTACGATTGCTGTTACAGGCGCTGTAAATTAGTTTACGCTGCT  
 CTGCAATCTGCAAGCAAAACGACTGGTTTCACTGTTGGTACAGTAGACACAGCTGGTTAT  
 ATCTCTGTAGGTACTGATGGGAGTGTTTCAAGGCATATGATGCTGCGACTTCTGGCAACAAA  
 GCTTCTTACCAACACTGACGGTACACTGACTACTGATAACACCACTAAACTGTATCTG  
 CAGAAAGATGGCTCTGTAAACCAACGGTTCAAGTAAAGCGGTCTATGTAGAAGCGGATGGT  
 GATTTCACTACCGAGCTGCAACCAAGCCGCAACCAACCCAGATCCGCTGGCGCGCTCTG  
 GATGACGCAATCAGCCAGATCGACAAGTTCCGTTATCCTTGGGTGCTATCCAGAACCGT  
 CTGGATTCTGCAGTCAACCACTGAACAACCAACCAACCACTGCTGTGAAGCGCAGTCC  
 CGTATTCAAGGACCGGACTATGCGACCGAAGTGTCCAATATGCGAAAGCGCAGATCATC  
 CAGCAGGCCGGTAACTCCGTGCTGGCAAAAGCCAAACAGGATCCGCGACGAGTTCTGTCT  
 CTGCTGCAGGGTTAA

Figure 25

48/96

AACAAATCTCAGTCTTCTCTGAGCTCCGCCATTGAA  
CGTCTCTCTTCTGGCCTGCGTATTAAACAGTGCTAAAGATGAGCGAGGTCAGGCGATT  
GCTAACCGTTTTACAGCAAATATTAAAGGCTGACTCAGGCTTCCCGTAACGCGAATGAT  
GGTATTTCTGTTGCGCAGACCACTGAAGTGCGCTGAATGAAATTAACAACAACCTGCAG  
CGTATTCTGGAACCTTTCTGTTCAAGCAACTAACGGTACTAACTCTGACAGCGATCTTTCT  
TCTATCCAGGCTGAAATTACTCAACGTCTGGAAGAAATGACCGTGTATCTGAGCAAACT  
CAGTTTAAACGCGGTGAAAGTCCTTGCTGAAATAATGAAATGAAATTCAGGTTGGTGCT  
AATGATGGTGAAACCATCACTATCAATCTGGCAAAAATGATGCGAAAACTCTCGGCCTG  
GACGGTTTTAATATCGATGGCGCGCAGAAAGCAACCGGCAGTGACCTGATTTCTAAATTT  
AAAGCGACAGGTACTGATAATTATCAAAATAACGGTACTGATAACTATACTGTTAATGTA  
GATAGTGGCGTAGTACAGGATAAAGATGGCAACAAGTTTATGTGAGTACTGCGGATGGT  
TCACTTACGACCAGCAGTGATACTCAATTCAGATTGATGCAACTAAGCTTGAGTGGCT  
GCTAAAGATTTAGCTCAAGGGAATAAGATTGTCTACGAAGGTATCGAATTTACAAATACC  
GGCACTGTGCTATAGATGCCAAAGGTAATGGTAAATTAACGCCAATGTTGATGGTAAG  
GCTGTTGAATTCACTATTTCTGGGGAGTACTGATACATCAGGTACTAGTGCAACCGTTGCC  
CCTACGACAGCCCTATACAAAATAGTGACGGGCAATTGACTGCAACAAAAGTTGAAAAT  
AAAGCAGCGACACTATCTGATCTTGATCTGAACGCTGCCAAGAAAAACAGGAAGCAGTTA  
GTTGTTAAACGGTGCAACTTACGATGTTAGTGCAGATGGTAAACGATAACGGAGACTGCT  
TCTGGTAAACAATAAAGTCATGTATCTGAGCAAAATCAGAAGGTGGTAGCCCGATTCTGGTA  
AACGAAGATGCGCAAAATCGTTGCAATCTACCAACCAACCGCTCGAAACTATCGACAAA  
GCATTGGCTAAAGTTGACAATCTGCGTTCTGACCTCGGTGCAGTACAAAACCGTTTCGAC  
TCTGCCATACCAACCTTGGCAACACCGTAAACAACCTGTCTTCTGCCCGTAGCCGTATC  
GAAGATGCTGACTACGCGACCGAAGTGCTAACATGTCTCGTGGCAGATCCTGCAACAA  
GCGGGTACCTCTGTTCTGGCACAG

Figure 26

49/96

ATGGCACAAGTCATTAATACCAACAGCCTCTCGCTGATCACTCAAAATA  
 ATATCAACAAGAACAGTCTGCGCTGCGAGTTCTATCGAGCGTCTGTCTTCTGGCTTGC  
 GTATTAACAGCGCGAAGGATGACGCGAGCGGTGAGGCGATTGCTAACCGTTTTACTTCTA  
 ACATTAAAGGCCTGACTCAGGCGGCACGTACGCCAACGACGGTATCTCTTGGCGCGAGA  
 CCACCGAAGGTGCGCTGTCTGAAATCAACAACACTTACAGCGGTACGTGAACGTGACCG  
 TTCAGGCAACACCGGTACTTAATCCGACTCCGACCTGGCTTCTATTACGACGAAATCA  
 AATCCCGTCTGGATGAAATTGACCGGTATCTGGTCAGACTCAGTTCAACGGCGGTGAACG  
 TGCTGGCAAAAGACGGTTCATGAAAATTGAGGTAGGTGCTAACGACGCCAGACTATCA  
 CTATTGACCTGAAAAAAATCGACTCTGATACTCTGGGCGCTGAATGGTTTTAACGTGAATG  
 GTTCTGGGACGATTACCAACAAGCAGCAACTGTCAGTGATGTTACTCGCGCAGGCGGTA  
 CATTGGTGAATGGTGCCCTATGATATAAAAACCACTAACACAGCGCTGACTACAACGTGATG  
 CCTTCGCGAAATGAATGATGGTGATGTTGTTACTATCAATAATGGTAAGGATACTGCCT  
 ATAAATATAATGCTGCTACAGGTGGGTTACGACGGATGTCCTCATCTCCGGGGATCCTA  
 CCGCTGCTGACGCTACTGCTAATAAACTGCCCGTGATGCACTTCGGCGCTCTTTACATG  
 CTGAGCCGGGTAATACTGTTAATGGTCTTGGACTACGAATGATGGTACGGTAAAAATTG  
 ATACCGATGCCGATGGTAAGATTCTATTGGTGGTGTGGTCTTATGTAGATGCAGCAG  
 GCAACCTGACCACTAACGCGAGCAGGTATGACGACTCAAGCAACAACACCGATTTGGTTA  
 CTGCTGCTGCATCTGCTACTGTGTAAGGGTGGATCCCTGACCTTGGTGACACGACGTATA  
 AAATTGGTCAGGCTACGGCTGGGTTGATCCTGATGACGCTTCAGATGATGTACTGGGCA  
 CCATTTCTTACTCTAAATCAGTAAGCAAGGATGTGTTCTTGCTGATACTAAGCAACTG  
 GTAACACGACAACAGTTGATTTCAACTCCGGTATCATGACTTCAAAGGTTAGTTTCGATG  
 CAGGTACATCAACTGATACATTCAAAGATGCAGATGGTGCTATCACCRAAACTAAAGAA  
 ACACCACTTCTTATGCTGTAATAAAGATACTGGTGAAGTTACCGTTGCTGATTATGCTG  
 CGGTAGATAGCGCGGATAAGGCTGTTGATGATACTAAATATAAACCGACTATCGGCGCGA  
 CAGTTAACTGAATTCGCGAGTAAATTGACCACTGATACCACAGTGCAGGCACAGCAA  
 CCAAGATCCTCTGGCTGCCCTGGACGCTGCTATCAGCTCCATCGACAAATTCGGTTCAT  
 CCTGGGTGCTATCCAGAACCGTCTGGATTCCGCGAGTCACCAACTGAAACAACACCACTA  
 CCAACTGTCCGAAGCGCAGTCCCGTATTCAGGACGCCGACTATGCGACCGAAGTGTCCA  
 ACATGTCCGAAGCGCAGATTATCCAGCAGGCGCGTAACCTCCGTGCTGGCAAAAGCCAAAC  
 AGGTACCGCAGCAGGTCTGTCTCTGCTACAGGGTTAA

Figure 27



50/96

AACAAAAACAGTCTGCGCTGTCGACTTCTATC  
 GAGCGCCTTCTTCTGCTGCGTATTAAACAGCGCTAAAGATGACGCTGCGGGCCAGGCG  
 ATTGCTAACCGCTTCACTTCTAACATCAAAGGTCTGACTCAGGCCGACGTAACGCCAAC  
 GACGGTATTTCTCTGGCGCAGACCACTGAAGCGCGCTGTCTGAGATTAACAACAACCTTG  
 CAGCGTGTGCGTGAGTTGACTGTACAGGCGACGACCGGACTAACCTGTATCTGACCTG  
 TCTTCTATCCAGGATGAAATCAAATCCCCTTTAAGCGAAATTGACCGTGTATCTGGTCAG  
 ACTCAGTTTAAACGGGTGAACGTACTGGCTAAGAATGACACCCTGTCTATTCAAGTAGGT  
 GCAAATGACGGTCAGACTATCAATATTGACTGCAGCAAATCGATTCTCATACACTGSGT  
 CTGGATGGTTTCAGCGTTAAAAATAATGATGCAGTGAACCAAGTCTGCCGTGAATACT  
 CTTGGGGGGGGGGCAGGTTCTGTTGCTGTCGACTTCGCAACAACCAAGTTTGACTGTATC  
 ACTGCTCTCGGTAGCGGTGCTATCAGCGAAATTGCTAAAGACGATAATGGTGATTACTAC  
 GCGCATGTACAGGGACTACGGTAATACTGCTGATGGTTACTATGCTGTGATATCGAC  
 AAGGCTACCGGTGAGGTGCTCTGAAAGATGGTAACGTAGATACACCGACAGGTACGCCA  
 ACGACGACAAGCACATATGACTTCACAGACGCTGGTCAAACCGTTTCCTTTGGCACTGAT  
 GCTGCAACAGCCGTTATCAGCACTGGTGCTTCTCTGTTAAACTTCAGGATGAGAAAGGC  
 AATGATACTGCTACTTATGCAATCAAAGCACAAGATGGCAGCCTGTATGCCGCCAACGTT  
 GATGAGGCTACCGGTAAAGTCACTGTCAAAACCGCCAGCTATACTGATGCTGACGGCAAA  
 GCAGTGACCGATGCCGCTGTAAACTGGGTGGTGACAATGGCACAACCGAAATTGTTGTC  
 GATGCTGCGTCAGGTAAACTTACGATGCTGGTGCACTGCAAAACGTTGATCTCTCCAGT  
 GCAACCAACACGGTTAACCGCAATCCGGAACGGTAAACACAGCTCTCCGCTGGCTGCCCTT  
 GACGACGCAATCAGCCAGATCGCAAAATTCGCTCTCCCTCCGTTGCGGTGCAGAACGCT  
 CTGGATTCCGCGGTACCAACCTGAACAACACCACTACCAACCTGTCTGAAGCGCAGTCC  
 CGTATTGAGGACGCTGACTATGCGACGGAAGTATCCAACATGTCGAAAGCGCAGATCATC  
 CAGCAGGCAGGTAATCCGTGCTGTCAAA

Figure 28

GCCTGTGCGACTTCTATCGAGCGCTCTCTTCTGGTCTGCGCATTAAACAGCGCTAAAG  
ATGACCGCTGCGGGCCAAAGCGATTGCTAACCGCTTCACTTCTAACATCAAAGGTCTGACTC  
AGGCCGACGTAACGCGCAACGACGGTATTTCTCTGGCGCAGACCACTGAAGGCGCACTGT  
CTGAAATCAACAACAACCTTGACGGTGTTCGTGAACGACCGTTACGGCCACTACCGGTA  
CTAACTCTGATTCTGACCTGTCTTCAATACAGGACGAAATCAAATCCCGTCTCGATGAAA  
TTGACCGCGTATCCGGTCAGACTCAGTTCAACGGCGTTAATGTTCTTTCCAAAGATGGTT  
CAATGAAAATTACGGTTGGTGCGAATGATGGTCAAACCTATCTCCATCGATCTGAAGAAAA  
TTGATTCTTCAACTTTGGGGCTGAATGGCTTCTCAGTTTCTAAAACTCTCTTAATGTCA  
GCAATGCTATCACATCTATCCCGCAAGCGCTAGCAATGAACCTGTTGATGTTAACTTCG  
GTGATACTGATGAGTCTGCAGCAATCGCAGCAAATTGGGGGTTTCGATACGTCAAGCC  
TGTGCTGCACAACATCCTTGATAAAGATGGTAAGGCAACAGCTGATTATGTTGTCAGT  
CAGGTAAGAGACTTCTATGCTGCTTCTGTTAATGCCGCTTACAGTAAAGTAACCTTAAACA  
CCATTGATGTTACTTATGATGATTATGCGAACGGTGTGACGATGCCAAGCAACAGGTC  
AGCTGATCAAAGTTTCAGCAGATAAAGACGGCGCAGCTCAAGGTTTGTCACTTCAAG  
GCAAAAATATTTCTGCTGGTGATGCGGCAGACATTCTTAAGAATGGAGCAACAGCTCTTA  
AGTTAACTGATCTGAATTTAAGTGATGTTACTGATACTAATGGTAAGGTAACCACAACCTG  
CGACTGAGCAATTTGAAGTGCTTCACTGAGGATCCGCTGGCGCTTCTGGATAAAGCTA  
TTGCATCAGTCGACAAATTCGGTCTTCTCTAGGTCCGTGCGAAGCCGTCTCGATTCCG  
CTATCACCACCTGAACAACACCACCAACCTCTGTAAGCGCAGTCCCGTATTCCAG  
ACGCCGACTATGCGACCGAAGTGTCCAACATGTCGAAAGCGCAGATCATCCAGCAGGCA

Figure 29

52/96

ATGGCACAAAGTCATTAATACCAACAGCCTCTCG  
 CTGATCACTCAAAATAATATCAACAAGAACCAGTCTGCGTGTGCGAGTTCTATCGAGCGT  
 CTGTCTTCTGGCTTGGCTATTAACAGCGCGAAGGATGACGCCGAGGTCAGGCGGATTGCT  
 AACCGTTTTACTTCTAACATTAAAGGCTGACTCAGGCTGCACGTAACGCCAACGCGGT  
 ATTTCTGTTGCAAGACCACTGAAGGCGCGTGTCCGAAATCAACAACAACTTACAGCGT  
 ATTCTGGAACAGCGGTTACAGGCCACTACAGGACTAACTCCGATTCTGACCTGGACTCC  
 ATCCAGGACGAAATCAAAATCTCGTCTGGACGAAATGACCGCGTATCTGGTCAGACCCAG  
 TTCAACGCGGTGAACGTGCTGTCTAAAGATGGCTCGATGAAATTCAGGTCGGCGCGAAC  
 GATGGCGAAACGATTACTATTGATCTGAAGAAAATGACTCTGATACGCTAAATCTGGCT  
 GGTTTTAACGTAATGGTGTGGCTCTGTTGATAATGCCAAGGCGACTGGCAAGATCTT  
 ACTGATGCTGGTTTTACGGCAAGCGCAGCTGATGCTAATGGCAAAATCACTTATACCAA  
 GACACCGTTACTAAATTCGACAAAGCGACAGCGGCTGATGTATTGGGCAAGCGGCTGCT  
 GGCATAGCATTACCTATGCGGGCACTGATACTGGCTTAGGAGTCGCTGCTGATGCCTCG  
 ACTTACACCTACAATGCAGCCAAATAGTCTTACACTTTTGATGCTACTGGTGTGCGAAG  
 GCGGATGCTGGAACGGCACTGAAAGGGTACTTAGGCGCATCTAACACCGGTAAATTAAT  
 ATCGGTGGTACCGAGCAAGAAATTAACATTGCCAAAGATGGCTCCATCACCGATACCAAT  
 GCGGATGCGCTGTATCTCGATAGTACCGGCACTTAACCAAAAATACCGCAATTTGGGG  
 GCTGCTGATAAGCAACTGTAGATAAACTGTTGCTGGTCTCAGGATGCAACGATCACC  
 TTCGATAGCGCATGACAGCTAAATTCGATCAAACTGCTGGTACCGTTGATTTCAAAGGC  
 GCGTCTATTTCTGCTGATGCAATGGCATCAACCTTAAATTAATGGTTCCTATACAGCCAAC  
 GTAGGTGGTAAGGCTTATGCCGTAAACCGTGGCGCAGTTCAGACAGGTGGCGAGATGTG  
 TATAAAGATACCACTGGCGCACTGACGACTGAAGATGACGAAACCGTTACCGCGACCTAC  
 TACGGTTTTGCTGATGGTAAAGTTTCTGACGGTGAAGGTTCTACTGTCTATAAAGCTGCT  
 GATGGTTCATCACTAAAGATGCGACTACCAAGTCTGAAGCAACCACTGACCTCTGAAA  
 GCCCTTGACGACGCAATCAGCCAGATCGACAAATTCGCTCTCTCCCTCGGTGCCGTTCAA  
 AACCGTCTGGATTCCGCCGTCAACCACTGAACAACCACTACCAACCTGTCTGAAGCG  
 CAGTCCCGTATTACAGGACGCCGACTATGCGACGAAAGTGTCAACATGTCGAAGCGCAG  
 ATCATTACGACGCGGTAACCTCCGTGCTGGCAAAAGCCAACCGGTACCGCAGCAGGTT  
 CTGTCTCTGCTGCAGGGTTAA

Figure 30

53/96

AACAAATCTCAGTCTTCTCTTAGCTCTGCTATTGA  
 GCGTCTCTCTTCTGGCCTGCGTATTAACAGTGTCTAAAGATGACGCGAGGTCAGGCGAT  
 TGCTAACCGTTTACGGCAAATATTAAGGTCTGACTCAGGCTTCCGTAACGCGAATGA  
 TGGTATTCTGTGCGCAGACTACTGAAGGTGCGCTGAATGAAATTAACAACAACCTGCA  
 GCGTGATCGTGAACCTGACTGTTCAAGGCAACTAACGGTACTAACTCTGACAGCGATCTTTC  
 TTCTATTCAAGCAGAAATTAACAACGTCTGGAAGAAATGACCGGTATCTGAGCAAAAC  
 TCAGTTTAAACGGCGTGAAGTCTTGGCCGAAATAATGAAATGAAATTCAGGTTGGTGTC  
 TAATGATGGGGAACCATCACTATCAATCTGGCAAAATTTGATGCGAAACTCTCGGCCT  
 GGAACGCTTTAATATCGATGGCGCGCAGAAAGCAACTGGCAGTGACCTGATTTCTAAAT  
 TAAAGCGACAGGTACTGATAATTATCAAAATTAACGGTACTGATAACTATACTGTTAATGT  
 AGATAGTGGAGCAGTTCAAATGAGGATGGTGACGCAATTTTGTAGCGCTACCGATGG  
 TTCTCTGACTACTAAGAGTGATACAAAAGTCGGTGGTACAGGTATTGATGCGACTGGGCT  
 TGCAAAAGCCGCGAGTTTCTTTAGCTAAAGATGCCTCAATTAATACCAAGGTATTACTTT  
 CACCAACAAAGGCACTGATGCAITTTGATGGCAGTGGTAACGGCACTCTAACCGCTAATAT  
 TGATGGCAAAGATGTAACCTTTACTATTGATGCGACAGGGAAGGACGCAACATTA AAAAC  
 GTCTGATCCTGTTTACAAAAATAGTGCAAGTCAGTTCACTACAACCTAAGGTTGAAAACAA  
 AGCCGCTACAGCATCGGATCTGGACTTAAATAACGCTAAAAAAGTGGGTAGTTCTTTAGT  
 TGTAATGGCGCTGATTATGAAGTTAGCGCTGATGTTAAGACAGTAACCTGGGCTTGSCAA  
 AACTATGTATCTGAGCAATCAGAAGGTGGTAGCCCGATTCTGGTAAAGAAGATGCAGC  
 AAAATCGTTGCAATCTACTACCAACCCGCTCGAAACCATCGACAAGGCATTGGCTAAAGT  
 TGACAATCTGCGTTCTGACCTCGGTGCAAGTACAAAACCGTTTCGACTCTGCTATCACCAA  
 CCTTGGCAACACCGTAAACAACCTGTCTTCTGCCGTAGCCGTATCGAAGATGCTGACTA  
 CGCGACCGAAGTGTCTAACATGTCTCGTGGCAGATCCTGCAACAAGCGGGTACCTCTGT TCTGGCGCAG

Figure 31

54/96

ATGGCACAAGTCATTAATACCAACAGCCTCTCGGTGATCACTCAAAA  
 ATATCAACAAGAACAGTCTGCGCTGTCGAGTTCTATCGAGCGTCTGTCTTCTGCGTTGC  
 GTATTAAACAGCGCAAGGATGACGCCGAGGTGAGCGATTGCTAACCGTTTACTTCTA  
 ACATTAAGGCGCTGACTCAGGCTGCACGTAACGCCAAGATGGTATTTCTGTGTCACAGA  
 CCACTGAAGGCGCGCTGTCCGAATCAACAACAACCTTACAGCGTATCCGTGAAGTACCG  
 TTCAGGCTTCTACCGGACTAACTCCGATTGGATCTGAGTCCATTGAGGACGAATCA  
 AATCCGCTCTGGACGAATTAACCGCGTATCTGGCCAGACCCAGTTCAACGGCGTGAACG  
 TACTGGCGAAGACCGTTCAATGAAAATTCAGGTTGGTGCGAATGACGCCAGACTATCA  
 CGATTGATCTGAAGAAAATTGACTCTGATACGCTGGGGCTGAGTGGGTTAATGTGAATG  
 GTAGCGGGCTGTGGCTAATACTGCAGCGACTAAATCTGATTGGCAGCAGCTCAACTCT  
 TGGCTCCAGGTAAGTCTGATGCTAATGGTACAGTTACCTATACTGTTGGCGCAGGCTGA  
 AAACATCTACAGCTGCAGATGTAATTGGGAGTTTGGCTAATAACGCAAAAGTTAATGCCA  
 CAATTGCAAAATGGTTTGGATCGCCACAGCTACAGATTATACATACAACAGCGCTACAG  
 GCGATTTTACATATAGTGCAACTATTGCAGCTGGTACAAAATCTGGTGATAGTAACAGTG  
 CTCAGTTACAATCCTTCTGACACCAAAAGCGGGCGATACTGTCAACTTAAACGTTAAAA  
 TTGGTTCTACGTCGAATTGACGTTGATTGGCTAGCGACGGTAAAATTACCGCAAAAGATG  
 GTTCAGAACTATTATTGACGTAGATGGTACCTCACTCAAAACAATGCTGGGACTGTCA  
 AAGCAGCCACTCTTGATGCACTGACTAAAAACTGGCATAACAACGCGACACCGAGTCCCG  
 TATCTACGGTAATTACAACCTGAAGATGAAACAACCTTCACTCTGGCTGGCGTACTGATG  
 CTACTACTTCTGGTGCAATCACTGTAGCAAAATGCAAGAATGAGTGCTGAGTCTCTTCAAT  
 CGGCAACTAAGTCCACAGGATTCACAGTTGATGTTGGAGCTACTGGTACCAGCGCAGGCG  
 ATATTAAAGTTGATAGTAAAGGTATAGTACAACAACACAGGTACAGGTTTGAAGACG  
 CTTACACCAAAAGCTGATGGTTCACTGACTACCGATAATAACAACCAATCTGTTTTTGCAAA  
 AAGACGGAAGTGTGACCAATGGTTCAAGGTAAGCAGTCTATGTTTCAGCGGATGGTAATT  
 TTACTACTGACGCTGAAACTAAAGCTGCAACACCGCCGATCCACTGAAAGCTCTGGACG  
 AAGCGATCAGCTCCATCGACAAATTCGGTCTCTCCCTCGGTGGCGTGCAAAACCGTCTGG  
 ATTCGCGAGTCACCAACCTGAACAACACCACTACTAAGCTGTCTGAAGCGCAGTCCCGTA  
 TTCAGGACGCTGACTATGCGACCGAAGTGTCCAATATGTGCAAAACGCGAGATCATCCAGC  
 AGGCCGCTAACTCCGTGTGGCAAAAGCTAACAGGTACCGCAGCAGGTTCTGTCTGCTG TGCAGGGTTAA

Figure 32

55/96

AACAAAAACCAGTCTGCGCTGTCGACTTCTATCGAGCGCCTCTCTT  
CTGGTCTGCGCATTAAACAGCGCTAAAGATGACGCTGCGGGCCAGGCGATTGCTAACCGCT  
TCACTTCTAACATCAAAGGTCTGACTCAGGCCGCGACGTAACGCCAACGACGGTATCTCTC  
TGGCGCAGACCACTGAAGGCGCACTGTCTGAAATCAACAACAACCTGACGCGTGTTCTG  
AGCTGACCGTTTCAGGCCACTACCGGTACTAACTCTGATTCTGACCTGTCTTCAATCCAGG  
ACGAAATCAAATCCCGTCTCGATGAAATTGACCGGTATCCGGTCAGACTCAGTTCAAAG  
GCGTGAACGTACTGGCAAAAGATAACACCATGAAGATTCAAGTTGGTGGCAACGATGGTC  
AGACTATATCCATCGACCTGCAAAAAATCGACTCTTCTACTCTTGGTTTGAACGGTTTCT  
CCGTTTCTAAAAATGCTCTGAAACTAGCGAAGCGATCACTCAGTTGCCGAACGGTGCGA  
ATGCACCAATCGCTGTGAAGATGGATGCGTCTGTTCTGACCGATCTTAACATTACTGATG  
CTTCGCTGTTTCGCTGCAACGTAACCTAAAGGTGGTGTGCGCAACGCTCACTTATGTTG  
TTCAGTATGGCGATAAGAGCTATGCAGCATCTGTTGATGCGGGAGGTACAGTAAACTGA  
ATAAGCCGACGTAACATATAACGACGCGAGCAAAATGGTGTACGAATGCCACCCAGATTG  
GTAGTCTGGTTTCAGGTTGGTGCTGATGCAACAATGATGCAGTTGGTTTTGTTACCGTGC  
AGGGGAAAACTATGTTGCTAATGACTCATTAGTCAATGCTAATGGCGCTGCTGGCGCTG  
CAGCAACTAGAGTTACAATTGATGGTGATGGTAGCCTTGGAGCTAACCCAGGCTAAAAATG  
AACTTAGCCAAAAATGGTGCTACTGCTGCAACATCAGAGTTCGCTGGTGCTTCAACCAACG  
ATCCACTGACTCTGCTGGACAAAGCTATCGCATCTGTTGATAAATTCCGTTCTTCTTTGG  
GGGCGGTACAGAACCGTCTGAGCTCCGCTGTAAACCAACCTGAACCAACACCACTACCAACC  
TGCTGAAGCGCAGTCCCGTATTCAGGACCGCGACTATGCGACCGAAGTGCCAAACATGT  
CGAAAGCGCAGATCATCCAGCAGGCAGGTAACCTCCGTGCTGTCCAAA

Figure 33

56/96

ATGGCACAAGTCATTAATACCAACAGCCTCTCGCTGATCACTCAAATAATATCAACAAGA  
 ACCAGTCTGCGCTGTCGAGTTCTATCGAGCGTCTGTCTTCTGGCTTGCATTAAACAGCG  
 CGAAGGATGACGCCGAGGTTCAGGCGATTGTCTAACCGTTTACTTCTAACATTAAAGGCC  
 TGA CT CAGGCTGCACGTAACGCCAACGACGGTATTTCTGTTGCACAGACCACTGAAGGCG  
 CGCTGTCCGAAATCAACAACAATTACAGCGTATTTCGTGA CT GACGGTT CAGGCGACGA  
 CCGGAAC TAACTCCACCTCTGACCTGGACTCCATT CAGGACGAAATCAAATCCCGCTCTTG  
 ATGAAATTGACCGCGTATCCGGCCAAACCCAGTTCAACCGCGTGAACGTACTGTCAAAAG  
 ATGGCTCGATGAAAATTCAGGTCCGGCGCAAATGATGGTGA AACCATCACGATTGATCTGA  
 AAAAGATCGACTCTTCTACATTGAAGCTGACCAGCTTCAATGTTAACGGTAAAGGCGCTG  
 TTGATAATGCTAAAGCCACTGAAGCAGATCTGACCGCTGCGGGCTTCTCCAAGGTGCAG  
 TCGTCAGTGGCAACAGCACCTGGACTAAATCTACTGTTACTACCTTTAATGCAGCAACAG  
 CTACCGACGTGCTGGCAAGCGTTAGCGGCGGCAGCACTATTAGCGGTTATACCGGTACAA  
 ACAATGGATTAGGCGTAGCGGCTTCTACTGCATATACCTACAACGCAACCAGCAAGTCTT  
 ATTCAATTGACGCAACCGCACTTACCAATGGCGATGGTACTGGGGCCACCACTAAAGTTG  
 CTGATGTGCTGAAAGCCTATGCAGCAAAACGGTGATAATACGGCTCAGATCTCCATCGGCG  
 GAAGCGCTCAGGACGT TAAAAATGCCAGCGATGGCACCCCTGACTGACGTCAATGGTGATG  
 CTTTATATATTGGTCTGACGGCAACCTGACTAAAAACAGGCGCGGTCAGATGCGG  
 CAACGTTGGACGTTATTTCAACGGTGCGAATGGTAATGCAGCAGTTGATGCGAAGATTA  
 CATTCGGCAGCGCATGACCGTTGATTTTCAACCCAGGCTAGCAAAAAGTGGATATTAAAG  
 GCGCAACGGTATCCGCGAAGATATGGACACTGCGTTAACTGGGCAGGCTTATACCGTAG  
 CTAACGGCGCACAGTCTTTTGACGTTGCCGCTGGTGGGGCAGTAACCGCTACTACAGGTG  
 GCGCTACCGTAAATATTGGTGCTGATGGTGA ACTGACGACTGCGACCAACAAGACTGTCA  
 CAGAAACTTATCAGAAATTTGCTAACGGCAATATTTCTGGATGATGACGCGCGGCTCTGT  
 ACAAAAGCGGCTGACGGTTCTCTGACC ACTGAAGCTACTGGTAAATCCGAAGTGACCAAG  
 ATCCGCTGAAAGCGCTGGACGATGCTATCGCATCCGTAGACAAATCCGCTCCTCCTCG  
 GTGCGGTGCAGAACCGCTCGGATTTCCGCAGTCACCAACCTGAACAACCACTACCAACC  
 TGTCTGAAGCGCAGTCCCGCATT CAGGACGCGACTATGCGACCGAAGTGTCCAATATGT  
 CGAAAGCGCAGATCATCCAGCAGGCGGTA ACTCCGTGCTGGCAAAAGCCAACCAAGTAC  
 CGCAGCAGGTTCTGTCTCTGCTGCAGGGTTAA

Figure 34

57/96

ATGGCACAAGTCATTAATACCAACAGCCTCTCGCTGATCAC  
 TCAAAATAATATCAACAAGAACCAGTCTGCGCTGTGAGTTCTATCGAGCGTCTGTCTTC  
 TGGCTTGGCGTATTAACACGCGCTAAGGATGACGCGCGGGTCAGGCGATTGCTAACCGTTT  
 TACTTCTAACATTAAAGGCGCTGACTCAGGCTGCACGTAACGCCAACGCGGTATTTCTGT  
 TGGCGAGACCACTGAAGGCGCGCTGTCCGAAATCAACAACACTTACAGCGTATCCGTGA  
 ACTGACGGTTTACGGCTTCTACCGGGAATTAACCCGATTGCGATCTGGACTCCATTACAGGA  
 CGAAATCAAATCCCGTCTGACGAAATTAACCGCGTATCTGGCCAGACCGAGTTCAACGG  
 CGTGAACGTACTGGCGAAGACGGTTCAATGAAATTCAGGTTGGTGCAGATGACGCGCA  
 GACTATCACTATTGATCTGAAGAAATTAAGTCTCAGATACGCTGGGGCTGAGTGGGTTTAA  
 TGTGAATGGTGGCGGGCTGTGCTAATACTGCAGCGACTAAAGATGATTTGGTCTGCTGC  
 ATCAGTTTACGCTGCGGTAGGTAATGAATACACTGTCTCTGTGGCTGTGAAATCAAC  
 TGCTGCTGATGTTATTGCTAGTCTCACAGATGGTGGCAGTAAGTGGCGGTGGTGAAG  
 CAATGGTTTGGTGCAGGGGCAACTGGAGATGCTTATAAATTCATCAAGCAACCAACAC  
 TTTTACTTACAATACCACTCAACAGCGGCGAGAACTCCAATCTTACCTCAGCGCTAAGGC  
 GGGGGATACCGCAACTTTCTCCGTTGAAATTTGGTGGCACCAGCAGGATGTTGTTCTGGC  
 TAGTGATGGCAAAATCACAGCAAAAGACGGGTCTAAACTTTATATTGACACCACAGGGAA  
 TTTAACCCAAAACGGTGGAGGTACTTTAGAAGAAGCTACCCTCAATGGCTTAGCTTTCAA  
 CCACTCTGGTCCAGCGCTGCTGTACAATCTACTATTACTACTCGGATGGAACTTCAAT  
 AGTTCTAGCAGGTTCTGGCGACTTTGGAACAACAAAACCTGCTGGGGCTATTATGTACAC  
 AGGAGCAGTGATCAGTGTGATGCACCTTCTTCCGCCAGTAAAGCGACTGGGTTTACTTC  
 TGGCACTTATACCGTAGGTACAGATGGAGTTGTTAAATCTGGTGGCAATGACGTTTATAA  
 CAAAGCTGACGGGACGGGATTAACTACTGACAATACCACAAATATTATTACAGATGA  
 CGGGTCTGTAACATAAGTTCTGGTAAAGCTGTGTATGCTGATGCAACAGGAAAACTAAC  
 TACTGACGCTGAACTAAAGCCGAAACACCGCGGATCCCTGAAAGCTCTGGACGAAGC  
 GATCAGTCCATCGACAAATCCGTTCTTCCCTCGGTGGGTGCAAAACCGTCTGGATTTC  
 CGCGGTCAACCACTGAACAACCACTACCAACCTGTCCGAAAGCGAGTCCCGTATTCA  
 GGACCGCGACTATGCGACCGAAGTGTCCAACATGTGAAAGCGCAGATCATCCAGCAGGC  
 CGGTAACCTCGTGTGGCAAAAGCTAAACAGGTACCGCAGCAGGTTCTGTCTCTGCTGCA GGGTTAA

Figure 35



58/96

ATGGCACAAGTCATTATAACCAACAGCCTCTCGGTGATCAC  
 TCAAAAATAATATCAACAAGAACAGTCTGCGCTGTCGAGTTCTATCGAGCGTCTGTCTTC  
 TGGCTTGCGTATTAACAGCGCGAAGGATGACGCCGCGGGTCAGGCGGATTTGCTAACCGTTT  
 TACTTCTAACATTAAAGGCGCTGACTCAGGCTGCACGTAAACGCAACGACGGTATTTCCGT  
 TGGCGAGACCACCGAAGGCGCGCTGTCCGAAATCAACAACAATTACAGCGTATCCGTGA  
 ACTGACGGTTTCAAGCCACTACCGGTACTAACTCCGATTCTGACCTTGAAGTCCATCCAGGA  
 CGAAATCAAAATCTCGTCTTGATGAAATTGACCGCGTATCTGGTCAGACCCAGTTCAATGG  
 CGTGAATGTGTTGTCCAAAGACGGTTCAATGAAATTCAGGTGGGCGCAAATGATGGTGA  
 AACCATCAGGATTGACCTGAAAAAATCGACTCTTCTACACTGAAGCTGACCAAGCTTCAA  
 CGTCAACGGTAAAGGCGCTGTTGATAATGCAAAAGCCAAGCAAGCAGATCTGACCGCTGC  
 GGGCTTCTCCCAAAGTGCAAGTTGTGAGTGGCAATAGCACCTGGACTAAATCTACTGTTAC  
 TACCTTTAATGCAGCAACAGCTACCGATGTGCTGGCTAGCGTTAGTGGCGGCAGCACTAT  
 TAGCGGTTATGCTGGCACAACAATGGGTTAGGCGTAGCGGCTTCTACTGCATATACCTA  
 CAACGCAACAGCAAGTCTTATTCAATTGACGCAACCGCACTTACTAATGGTGATGGTAC  
 TGCGGGCTCAACTAAAGTTGCTGATGTTCTGAAAGCCTATGCAGCAACGCGCATAAACAC  
 GGCTCAGATCTCCATCGGTGGTAGCGCTCAGGAAGTTAAAAATTGCCAGCGATGGTACCCCT  
 GACGGTACTAATGGCGATGCTTTATACATTGGTGCTGACGGTAACTGACGAAAAACCA  
 GGCGCGCGGCCAGCGCGCGCAACGTTGGACGGTATTTTCAACGGTGCGAATGGTCATGA  
 TGCAGTTGATGCGAAGATTACCTTCGGCAGCGGCATGACCGTTGACTTCAACCCAGGTTAG  
 CAACAATGTGGATATTAAGGGCGCGACGGTATCCGCCGAAGATATGAACACTGCGTTAAC  
 CGGTACGGCTTATACCGTAGCTAACGGCGCACAGTCTTATGACGTTGGCGCTGATGGTGC  
 AGTAACTGCTACTACAGGTGGAGCGACCGTAAATATTGGTGCTGAGGGTGAACCTGACGAC  
 TGCGGCCAACAAAGACTGTCTACAGAACTTATCAAGAAATTTGCTAACGGCAATATTTCTGGA  
 TGATGACGGCGCGGCTCTGTATAAAGCGGCTGACGGCTCTCTGACCCTGAAGCTACAGG  
 TAAATCTGAAGCGACCAAGGATCCGCTGAAAGCGCTGGACGATGCTATCGCATCCGTAGA  
 CAAATTCGGTCTTCCCTGGGTGCGGTGCAGAACCGCTGGATTCCGCACTACCAACCT  
 GAACAACACCACTACCAACCTGTCCGAAGCGCAGTCCCGTATTTCAGGACGCCGATATGC  
 GACCGAAGTGTCCAACATGTCGAAAGCGCAGATTATTCAGCAGGCAGGTAACCTCCGTGCT  
 GGCAAAAGCTAACACAGTACCGCAGCAGGTTCTGCTCTGCTGCAGGGTTAA

Figure 36

59/96

AACAAAAACCAGTCTGCGTGTGACTTCTAT  
 CGAGCGCCTCTCTTCTGGTCTGCGCATTAAACAGCGCTAAAGATGACGCTGCGGGCCAGGC  
 GATTGCTAACCGCTTCACTTCTAACATCAAAGGTCTGACTCAGGCCGACGTAACGCCAA  
 CGACGGTATCTCTCTGGCGCAGACCACTGAAGGCGCACTGTCTGAAATCAACAACAACTT  
 GCAGCGTGTGCGTGAGTTGACTGTTGAGGCGACGACCGGGACTAACTCTGATTCTGACCT  
 GTCTTCTATTGAGGACGAAATCAAATCCCGTCTGGATGAAATTGACCGTGTTCGGGTCA  
 GACCCAGTTCAACGGCGTGAACGTGCTGGCTAAAAACGGTTCTATGGCGATTGAGTTGG  
 CGCGAATGATGGGCAGACCATCAACATCGACCTGCAGAAAAATCGACTTCTTACTCTGGG  
 CCTGGGCGGCTTCTCCGTATCTAACATGCACTGAACTGAGCGATTCTATCACTCAGGT  
 TGGTGCAGTGGTTCACTGGCAGATGTGAAACTGAGCTCTGTTGCCTCGGCTCTGGGTGT  
 AGACGCAAGCACTCTGACTCTGCACAACGTACAGACCCAGCTGGCGCAGCAACAGCTAA  
 CTATGTTGTCTCTTCTGGTCTGACAACTACTCAGTATCTGTTGAAGATAGCTCCGGTAC  
 AGTTACGCTGAACACCACTGATATAGGTTATACCGATACCGCTAATGGCGTTACTACCGG  
 TTCCATGACTGGTAAGTACGTTAAAGTTGGAGCTGATGCATTGGGTGCTGCTGTAGGTTA  
 TGTCACCGTACAGGGACAAAACCTCAAAGCTGATGCTGGCGCGCTGGTTAACTCCAAGAA  
 TGCTGCTGGTAGTCAGAAATGTTACTTCTGCAATTGGCGATATTGCTAATAAAGCGAATGC  
 TAACATTACACTGGAACCTCTTCTGCAGATCCACTGGCTCTGCTGGACAAGCTATCGC  
 ATCTGTTGATAAATTCGGTCTTCTCTAGGGGCGGTGCAGAACCGTCTGAGCTCTGCTGT  
 AACCAACCTGAACAACCACTACCAACCTGTCGGAAGCGCAGTCCCGTATTGAGGACGC  
 CGACTATGCGACCGAAGTGCCCAACATGTGAAAGCGCAGATCATCCAGCAGGCGGGTAA  
 CTCGCTGCTGTCTAAA

Figure 37

60/96

ATGGCACAAAGTCATTAATACCAACAGCCTCTCGCTGATCA  
 CTCAAAATAATATCAACAGAACAGTCTGCGCTGTCGAGTTCTATCAGAGCGTCTGTCTT  
 CTGGCTTGGGTATTAACAGCGCGAAGGATGACCGCGCGGTGAGGCGATTGCTAACCGTT  
 TTACTTCTAACATTAAAGGCTGACTCAGGCTGCACGTAACGCCAATGACGGTATTCTG  
 TTGCACAGACCACTGAAGGCGCGTGTCCGAAATCAACAACACTTACAGCGTATTCTGTG  
 AACTGACGGTTCAGGCTTCTACCGGACTAAGCTGATTCTGGATCTGGACTCCATTTCAGG  
 ACGAAATCAAAATCCCGTCTCGACGAAATGACCAGCTATCCGGTCAGACCCAGTTCAACG  
 GCGTGAACGTACTGGCAAAGACGGTTTCGATGAAAATTGAGTTGGTGCGAACGACGGCC  
 AGACTATCACTATTGATCTGAAGAAAATTGACTCTGATACGCTGGGGCTGAGTGGGTTTA  
 ACGTAAATGGTAGCGCAGATAAGGCAAGTGTGCGGCGCAGCTGACGGAATGGTTAAAG  
 ACGGATATATCAAAGGGTTAACTTTCATCTGACGGCAGCACTGCATATCTAAAACTACAG  
 CAAATACTGCAGCAAAAGGATCTGATATTCTTGGCGCGCTTAAGACTGGCGATAAAATTA  
 CCGCAACAGGTGCAAAATAGCCTTGTGATTAATGCGACATCGACAACCTTATATTTATAATG  
 CAACCAGCAATACCTTCTCCTATACGGCTGACGGTGTAAACCAACGAATGCTGCAGCAA  
 ATCTCATACCTGCAGCAGGAAAAACGACAGCTGCATCAGTTACTATTGGTGGGACAGCAC  
 AGAATGTAAATATTGATGATTCGGGCAATATTACTTCAAGTGATGGCGATCAACTTTATC  
 TGGATTCAACAGGTAACCTGACTAAAAACAGGCGCGGCAACCCGAAAAAGCAACCGTTT  
 CTGGGCTTCTCGGAAATACGGATGCGAAAGGTACTGCTGTTAAAAACAACCATCAAGACAG  
 AGGCTGTGTAAACAGTTACAGCTGAAGGTAATACAGGTACTGTAAAAATTGAAGGTGCTA  
 CTGTTTCAGCATCTGCATTACGGGCAATGCATATTTCGCGCAACCCGGTGGGAATACTT  
 ATGCTGTTGCCGCAAAATAACTACAAATGGTTTCTGGCGGGGGATGACTTAACCCAGG  
 ATGCTCAAACTGTTTCAACCTACTACTCGCAAGCCGATGGCAGCGTACGAAATAGCGCAG  
 GCAAGAAATCTATAAGACGCTGATGGTGTCTACAGCACAGAGAATAAAACATCGAAGA  
 CGTCCGATCCATTGGCTGCGCTTGACGACGCAATCAGTCCATCGACAAATTCGGTTCAT  
 CCTTGGGTGCTATCCAGAACCGTCTGGATTCCGCGGTCAACCACTGAACAAACCACTA  
 CCAACCTGTCCGAAGCGCAGTCCCGTATTGAGGACGCCGACTATGCGACCGAAGTGTCCA  
 ACATGTGCAAGCGCAGATCATCCAGCAGGCCGGTAACTCCGTGCTGGCAAAAGCTAAAC  
 AGGTACCGCAGCAGGTTCTGTCTCTGCTGCAGGGCTAA

Figure 38

61/96

AACAAATCTCAGTCTTCTCTGAGCTCCGCCATTGAACGTCTCTCTTCTGGCCTGCGTA  
 TTAACAGTGTAAAGATGACGACGAGGTGAGGCGATTGCTAACCGTTTACAGCAAAATA  
 TTAAGGTCTGACTCAGGCTTCCCGTAACGCGAATGATGGTATTTCTGTGCGCAGACCA  
 CTGAAGGTGCGCTGAATGAAATTAAACAACACCTGCGAGCTGTACGTGAACCTGACTGTT  
 AGGCAACTAACGGTACTAACTCTGACAGCGATCTTCTTCTATCCAGGCTGAAATTAATC  
 AACGTCTGGAAGAAATTGACCGGTATCTGAGCAAACCTCAGTTTAAACGCGTGAAGTCC  
 TTGCTGAAATAATGAAATGAAATTCAGGTTGGTGCTAATGATGGTGAACCATCACTA  
 TCAATCTGCGCAAAATTTGATGCGAAACTCTCGGCCTGGACGGTTTTAATATCGATGGCG  
 CGCAGAAAGCAACTGGCAGTGACCTGATTTCTAAATTTAAAGCGACAGGTACTGATAACT  
 ATGATGTTGGCGGTGATGCTTATACGTGTTAACGTAGATAGCGGAGCTGGGTAAAGTACC  
 AACTTATTGATAGTGTATTTATGTTACAGATAATGCCGATGACTTTGTCTAGCAGCTCCAC  
 CGATTTTGAGAACGACAGCGACTTCCGTCCGAGCCGTGCCAGGTGCTGCCCTCAGATTACG  
 GTTATGCCGCTCAATTGCTGCGTATATCGCTTGCTGATTACGTGCGAGCTTTCCCTTCAG  
 CGCGGATTACATACAGCGGCCAGCCATCCGTATCCATATCACCACGTCAAAGGGTGACAG  
 CAGGCTCATAAGACGCCCCAGCGTCGCCATAGTGCGTTTACCAGTAACGTGCGCAACAAC  
 CGTCTCTCGGAGCCTGTACATACGCGTAAAAACGCGAGCGCTGGCGCGATTAGCCCCGAC  
 ATAGTCCCACTGTTGCTGCCATTTCCGCGCAGACGATGACGTCACTGCCCGCTGTATGCG  
 CGAGGTTACCGACTGCGGCTGAGTTTTTAAGTGACGTAAAAATCGTGTGAGGCCAACG  
 CCCATAATGCGGCGAGTTGCCCGCATCCAAAGCCATTATGCGCATATCAATGATTTTC  
 TGGTGCGTACCGGGTTGAGAAGCGGTGTAAGTGAACCTGAGTTGCCATGTTTACCGGCAG  
 TGAGAGCAGAGATAGCGCTGATGTCGCGCGGTGCTTTTGCGGTTACGCCACACCCCGTCA  
 GTAGCTGAACAGGAGGGACAGCTGATAGAAAACAGAGCCACTGGAGCACTCAAAAACAC  
 CATCATACATAAATCAGTAAGTTGGCAGCATTACCGCGAGCTGTTAAAGATACTACAG  
 GGAATGATATTTTGTAGTGCAGCAGATGGTTCACTGACAACTAAATCTGACACAAAACA  
 TAGCTGGTACAGGGATTGATGCTACAGCACTCGCAGCAGCGCTAAGAAATAAGCACAGA  
 ATGATAAATTCAGTTTAAATGGAGTTGAATTCACAAACAACCTGACGCGGATGGCAATG  
 GGAATGGTGATATTTCTGCAGAAATTGATGGTAAGTCAGTGACATTTACTGTGACAGATG  
 CTGACAAAAAGCTTCTTTGATTACGAGTGAGACAGTTTACAAAAATAGCGCTGGCCTTT  
 ATACGACAAACCAAGTTGATAACAAGGCTGCCACACTTTCCGATCTTGATCTCAATGACAG  
 CTAAGAAAAACAGGAAGCAGTTAGTTGTTAACGSGTGCAACTTACGATGTTAGTGACAGATG  
 GTAAAACGATAACGGAGACTGCTTCTGGTAACAATAAAGTCATGTATCTGAGCAAACTCAG  
 AAGGTGGTACGCCGATTCTGGTAAACGAAGATGACGCAAAATCGTTGCAATCTACACCA  
 ACCCGCTGAAACTATCGACAAAGCATTTGGCTAAAGTTGACAACTCGCGTTCTGACCTCG  
 GTGCAGTACAAAACCGTTTGCAGCTCTGCTATCACCAACCTTGGCAACCCGTAAACAACC  
 TGTCTCTGCCCCGTAGCCGTATCGAAGATGCTGACTACGCGACCGAAGTGTCTAACATGT  
 CTCGTGCGCAGATCCTGCAACAAGCGGGTACCTCTGTTCTGGCGCAG

Figure 39

62/96

AACAAGAACCAGTCTGCGCTGTGCGAGTTCTATCGAGCGTCTGT  
CTTCTGGCTTGCCTATTAAACAGCGCGAAGGATGACGCCGAGGTTCAGGCGATTGCTAACC  
GTTTTACTTCTAACATTAAAGGCCTGACTCAGGCTGCACGTAAACGCCAACGACGCGTATT  
CTGTTGCGCAGACCCAGGAAGGCGCGCTGTCCGAAATCAACAACAACCTTACAGCGTGTGC  
GTGAATGACCGTTCAGGCAACCCAGGTACCAACTCCAGTCTGACCTGGACTCTATCC  
AGGACGAATAAATCCCGTCTGGACGAATTGACCGCGTATCCGGTCAGACCCAGTTCA  
ACGCGGTGAACGTACTGGCAAAAGACGGTCCATGAAATTCAGGTTGGCGCGAACGATG  
GCCAGACCATCACTATCGACCTGAAGAAGATTGACTCTTCTACGCTGAACTGACTGGTT  
TTAACGTGAATGGCAAAGCAGCGGTTGATAATGCTAAAGCGACGGATGCAATCTGACTA  
CCGCGGTTTTACACAAGGCGTTGTGGATTCAAATGGTAATAGTACTTGGACTAAATCAA  
CTACGACTAATTCGATGCGGCAACTGCAGTAAACGCTACTAGCAGCAGTTAAAGATGGCA  
GCACAATCAATTACACCGGTACTGGTAATGGTTTAGGGATTGCTGCAACAAGTGCTTATA  
CATATCAGCATAGCACTAAATCCTATACCTTTGATTCTACGGGGGCTGCAGTAGCTGGTG  
CCGCGTCCAGCCTGCAAGGTACTTTTGGTACAGATACGAATACTGCAAAAAATCACCATCG  
ATGGTTCTGCTCAAGAAGTAAACATCGCTAAAGATGGGAAAAATACTGATACTGATGGTA  
AAGCTTTATATATCGATTCCACTGGTAATTTGACTAAGAACCGCTCTGTACTTTAACTC  
AGGCAACATTGAATGATGCTCTTACTGGTGCTAATTCAGTTGATGATACAAGGATTGACT  
TCGATAGCGGCATGCTCTGTCAACCTTGATAAAGTGAACAGCACTGTAGATATCACTGGCG  
CATCTATTTACGCCGTGCAATGACTAATGAGTTGACAGGTAAAGCCCTATACCGTAGTAA  
ATGGTGCAGAATCTTACGCTGTAGCTACTAATAACACAGTAAAAACGACTGTCTGATGCTA  
AAAAATGTTTATGTTGATGCTAGTGGTAAATTAACCTACTGATGACAAAGCCACTGTTACAG  
AAACTTATCATGAATTTGCGAATGGCAATATCTATGATGATAAAGGCGCTGCTGTTTATG  
CGGCGCGGATGGTTCTCTGACTACAGAACTACAAGTAAATCAGAAGCTACAGCTAACCC  
CGCTGGCCGCTCTGGACGACGCAATCAGCCAGATCGACAAATTCGGTTTCATCCCTGGGTG  
CTATCCAGAACCCTGTGGATTCCGCACTCACCAACCTGAACAACCACTACCAATCTGT  
CTGAAGCGCAGTCCCGTATTTCAGGACGCGCACTATGCGACCGAAGTGCCAAATATGTCGA  
AAGGCGAGATCATCCAGCAGGCAGGCAACTCCGTGCTGCGCAAAA

Figure 40

64/96

AACAAAAACCAGTCTGCGCTGTCGACTTCTATCGAGCGCCTCTTCTGGTC  
 TGGCGATTAAACAGCGCTAAAGATGACGCTGCGGGCCAGGCGATTGCTAACCGCTTCACTT  
 CTAACATCAAAGGTCTGACTCAGGCTGCACGTAACGCCAATGACGGTATTCTCTAGCAC  
 AGACAGCGGAAGGCGCGCTGTCAGAGATTAAACAACAATTGACGCGTGTGCGTGAGTTGA  
 CCGTGCAGGCAACCACTGGTACCAACTCTGATTCCGATCTCTTCTTATTACGGATGAAA  
 TTAATCTCGTCTGGATGAAATTGACCGCGTCTCTGGTCAGACCAGTTTAAACGGCGTGA  
 ACGTACTGGCTAAAAACGGTTCTATGGCAATTCAGGTTGGCGCGAACGATGGCCAGACTA  
 TCTCTATCGACCTGCAGAAAAATAGACTCTTCTACTCTGGGTCTGAGCGGCTTCTCTGTTT  
 CTCAGAACTCCCTGAAACTGAGCGATTCTATCACTACGATCGGCAATACTACTGCTGCAT  
 CGAAGAACCTGGACCTGAGCGCAGTAGCAACTAAACTGGGCGTGAATGCAAGCACCCCTGA  
 GCCTGCACGAAGTTCAGGACTCTGCTGCTGACGGTACTGCTACCTTCGTTGTTTCTCTG  
 GCAGCGACAACATATGCTGTGTCTGTAGACGCGGCTCTGGTGCAGTTAACCTGAACACCA  
 CTGACGTCACCTATGATGACGCTACTAATGGTGTACTGGCGGACTCAGAACGGTCAGC  
 TGATCAAAGTAACTTCTGACGCCAACGGGTGCAGCTGTTGGTTACGTAACCAATTAGGGTA  
 AAAACTATCAGGCTGGTGCAGACCGGTGTTGACGTTCTGGCGAACAGCGGTGTTGCAGCTC  
 CAACTACAGCTGTTGATACCGGTACTCTGCAACTGAGCGGTACTGGTCAACTACTGAGC  
 TGAAGGTACTGCAACTCAGAACCACTGGCACTATTGGACRAAGCTATCGCTTCTGTTG  
 ATAAATTCGGTTCTTCTGCGGTGCGGTACAGAATCGTCTGAGCTCTGCTGTAAACCAACC  
 TGAATAACACCACCACTAACCTGTCTGAAGCGCAGTCCCGTATTGAGATGCCGACTATG  
 CGACCGAAGTGTCAAATATGTCTAAAGCGCAGATCGTTACGAGCGCCGTAAC

Figure 42

65/96

AACAAATCTCAGTCTTCTCTTAGCTCTGCTATTGAGCGTCTGCTCTT  
GGTCTGCGTATTAAACAGCGCAAAAGACGATGCAGCAGGTGAGCGGATTGCTAACCGTTTT  
ACGGCAAATATTAAAGGCTTGACCCAGGCTTCCCGTAACGCAAATGATGGTATTTCTGT  
GCGCAGACCACTGAAGGTGCGCTGAATGAAATTAACAACAACCTGCAGCGTATTCGTGAA  
CTTTCTGTTGAGGCAACTAACGGTACTAACTCTGACAGCGATCTTTCTTATCCAGGCT  
GAAATTACTCAACGTCTGGAAGAAATTGACCGTGATCTGAGCAAACCTCAGTTTAAACGGC  
GTGAAAGTCCTTGCTGAAAAATAGAAATGAAAAATTCAGGTTGGTGCTAATGATGGTGAA  
ACCATCACTATCAATCTGGCAAAAAATTGATGCGAAAACTCTCGGCCCTGGACGGTTTTAAT  
ATCGATGGCGCGCAGAAAGCAACAGGCAGTGACCTGATTCTAAATTTAAAGCGACAGGT  
ACTGATAATTATGATGTTGGCGGTAAAACTTATACCGTGAATGTGGAGACGGCGCGGTT  
AAGAATGATGCTAATAAAGATGTTTTTTGTAAGCGCAGCTGATGGATCGCTGACGACCAGT  
AGTGATACTAAAGTATCCGGTGAAAGTATTGATGCAACAGAAGTACGCAAACTTGCAATA  
AAATTAGCTGACAAAGGCTCCATTGAATACAAGGGCATTACATTACTAACAACTAGGC  
GCAGAGCTTGATGCTAATGGTAAGGTGTTTTGACCGCAAATATTGATGGTCAAGATGTT  
CAATTTACTATTGACAGTAATGCACCCACGGGTGCGGCGCAACAATACTACAGACACA  
GCTGTTTACAAAAACAGTGCAGGCCAGTTCCACCACTACAAAAAGTGAAAAATAAAGCCGCA  
ACACTCTCTGATCTGGATCTTAATGCAGCCAAGAAAAACAGGTAGCACTTTAGTTGTAAT  
GGCGCCACCTACAATGTCAGCGCAGATGGTAAAACGGTAACTGATACTACTCTGGTGCC  
CCTAAAGTGATGTATCTGAGCAAATCAGAAGGTGGTAGCCCGATTCTGGTAAACGAAGAT  
GCAGCAAAATCGTTGCAATCTACCACCAACCCGCTCGAACTATCGACAGGCATTGGCT  
AAAGTTGACAATCTGCGTTCTGACCTCGGTGCGGTACAAAACCGTTTCGACTCTGCCATC  
ACCAACCTTGGCAACACCGTAAACCAACCTGTCTTCTGCCGTAGCCGTATCGAAGATGCT  
GACTACGCGACCGAAGTGCTCAACATGTCTCGTGCGCAGATCTGCAACCAAGCGGTACC  
TCTGTTCTGGCGCAG

Figure 43

66/96

ATGGCACAAGTCATTAATACCAACAGCCTCTCGCTGATCACT  
 CAAAATAATATCAACAAGAACAGTCTCGGCTGTCTGAGTTCTATCGAGCGTCTGTCTTCT  
 GGCTTGCCTATTAAACAGCGCGAAGGATGACGCCGAGGTGAGGCGGATTGCTAACCGTTTC  
 ACCTCTAACATTAAGGCCCTGACTCAGGCTGACGCTAACGCCAACGACGGTATTTCTGTT  
 GCACAGACCACCGAAGGCCGCTGTCCGAAATCAACAACCTTACAGCGTATCCGTGAA  
 CTGACGGTTGAGGCTTCTACCGGGACTAAGTCTGATTGCGATCTGGACTCCATTGAGGAC  
 GAAATCAAATCCCGTCTGGACGAAATTGACCGGCTATCCGGCCAGACCCAGTTCAACGGC  
 GTGAACGTGCTGGCGAAAGACGTTCAATGAAAATTCAGGTTGGTGCGAATGACGGCCAG  
 ACTATCACTATTGATCTGAAGAAAATGACTCTGATACTCTGGGTTTGAGTGGATTTAAT  
 GTGAATGGCAAAGGGGCTGTGGCTAACGCAAAAGCGACCGAAGCAGATTTAACGGGGGCT  
 GGTTTCTCTCAAGGAGCGGTGGATACAAAGGGAATAGTACTTGGACAAAATCAACCACC  
 ACCAATTACTCAGCTGCAACAACCTGCTGACTTGTATCGACCATTAAAGGATGGCTCTACT  
 GTTACATATGACGGGACAGACACCGGATTAGGGGTCGACGACGAGGAAATATATCTTAT  
 GATGCGAACAGTAAATCTTATTCTTCAATGCCAATGGTCTACGGGCGCAAATACCGCA  
 ACTGCACTCAAAGTTACTTGGGACAGGTGCTAACACCGCTAAAAATTTCTATCGGTGGT  
 ACAGAGCAGGAAGTGAATATTGCCAAGATGGCACTATTACAGATACGAATGGTGATGCG  
 CTCCTATCTGGATATTACCGGCACTGACTAAGAACTATGCGGGTTACCACTCTGCAGCA  
 ACGCTGGATAACGTATTAGCTTCCGCAACTGTAATGCCACTATCAAGTTTGATAGCGGT  
 ATGACGGTTGATTACACTGCAGGTACTGGCGCAATATTACAGGTGCATCCATTTCTGCA  
 ATGATGACATGGCCGCAAACTGAGCGGAAAGGCGTACACTGTTGCCAATGGTGTGAGTCT  
 TATGACGTTGCTGCAGTTACGGGGGCTGTAACAACACAGCAGGTAATTCACCTGTGTAT  
 GCCGATGCAGACGTAATTAACGACGAGTGCCAGTAATACGGTTACTCAGACTTATCAC  
 GAGTTTGTCTAATGGTAAACATTTATGATGACAAAGGCTCGTCACTGTATAAAGCTGCAGAT  
 GGCTCTCTGACTTCTGAAGCTAAAGGGAATCTGAAGCAACCGCCGATCCCCGTAAGGCT  
 CTGGACGAAGCCATCAGCTCCATCGACAAAATCCGCTCTCCCTCGGTGCGTTCAAAC  
 CGTCTGGATTCTGCGGTGACCACTGAACAACACCACTACCAACCTGTCTGAAGCGCAG  
 TCCCGTATTGAGGACGCCGACTATGCGACCGAAGTGTCAAATATGCGAAAGCGCAGATC  
 ATCCAGCAGGCGGTAATCCGTTGTGGCAAAAGCTAACAGGATACCGCAGCAGGTTCTG  
 TCTCTGCTGACGGTTAA

Figure 44



67/96

CGCGTGTGCACTTCTATCGAGCGCCTCTCTTCTGGTTTGCGCATTAACAGCGCTA  
AAGATGACGCTGCGGGCCAGGCGATTGCTAACCGCTTCACTTCTAACATCAAGGTCTGA  
CTCAGGCCGACGTAACGCCAACGACGCGTATCTCTCTGGCGCAGACCACTGAAGCGCAC  
TGTCTGAAATCAACAACAACTTGACGCGTGTTCGTGAACGTGACCGTTTCAGGCCACTACCG  
GTACTAACTCTGATTCTGACCTGTCTTCAATCCAGGACGAAATCAAATCCCCGCTTGGCTG  
AAATCGATCGTGTCTCTGGTCAGACCCAGTTCAACGGCGTGAACGTGCTGGCTAAAAACG  
GTTCTCTGAATATTACAGGTTGGCGCGAATGATGGGCAGACCATCTCTATCGATTTCGAGA  
AAATAGACTCTTCTGCCCTTGGTTTAAAGTGGTTTTAGTGTTCGCGGTGGGGCGCTAAAAAT  
TAAGCGATACAGTGACGCGAGTCCGCGATGGTTTCAGCCGCGCCAGTTAAAGTGGATCTGG  
ATGCAGCAGCAACAGATATTGGTACTGCTTTGGGGCAAAAGGTTAATGCAAGTCTCTTAA  
CGTTGCACAATATCTTAGACAAAGATGGTGCGGCAACTGAGAACTATGTTGTTAGCTATG  
GTAGTGATAATTACGCTGCATCTGTTGCAGATGACGGGACTGTAACCTCTTAATAAAACGG  
ATATTACTTATTACGGCGGTGATATTACCGCGCTACCAAAGATGATACGTTGATTAAAG  
TTGCTGCTAATTCTGACGAGAGGCCGTTGGTTTCGCTACCGTTTCAGGGTAAGAATTATG  
AAATTACAGATGGTGTAAAAAACAGTCCACTGCTGCACCAACCGATATTGCTCAGACCA  
TTGATCTGGATACGGCTGATGAATTTACTGGGGCTTCCACTGCTGATCCACTGGCACTTT  
TAGACAAAGCTATTGCACAGGTTGATACTTTCCGCTCTCCCTCGGTGCCGTTCAAAACC  
GTCTGGATTCCGCACTCACCACCTGAACAACACTACTACCAACCTGTCTGAAGCGCAGT  
CCCCTATTTCAGGACGCCGACTATGCGACCGAAGTGTCCAATATGTCGAAAGCGCAGATCA TCCAGCAGGCC

Figure 45

68/96

ATGGCACAAAGTCATTAATACCAACAGCCTCTCGCTGATCACT  
CAAAATAATATCAACAAGAACCAAGTCTGCGCTGTCGAGTTCTATCGAGCGTCTGTCTTCT  
GGCTTGGCTATTAAACAGCGCAAGGATGACGACGCGGTGAGGCGATTGCTAACCGTTT  
ACTTCTAATATTAAGGCCTGACTCAGGCTGCACGTAACGCCAATGACGGTATTCTCTG  
GCGCAGACCACTGAAGGCGCACTGTCTGAAATCAACAACACTTGACGGTGTGCGTGAA  
CTGACCGTACAGGCGCAACCGGAACGAACTCCGAATCTGACCTGTCTCTATCCAGGAC  
GAAATCAAATCCCGTCTGGAAGAGATTGACCGCGTATCCGGCCAGACTCAGTTCAACGGC  
GTGAATGTGCTGGCAAAAGACGGCACCATGAAAATTGAGGTAGGCGCGAACGATGGTCAG  
ACTATCTCTATCGATCTGAAAAAATCGACTCTTCAACCTCGGGCTGACCGGTTTGTAT  
GTTTCGACGAAAGCGAATATTTCTACGACAGCAGTAACGGGGCGGCAACGACCACTTAT  
GCTGATAGCGCGGTTGCAATTGATATCGGAACGGATATTAGCGGTATTGCTGCTGATGCT  
GCGTTAGGAACGATCAATTTGATAATACAAACAGGCAAGTACTACGCACAGATTACCAAGT  
GCGGCCAATCCGGGCCTTGATGGTCTTATGAAATCCATGTTAATGACGCGGATGGTTCC  
TTCACGTAGCAGCGAGTGATAAAACAGCGGGTGCTGCTCCGGGTAAGTCTCTGACAAAGC  
GGTAAAGTTCAGACTGCAACCAACGACGCGAGGTACGGCTGTTGATGTCACTGCGGCTAAA  
ACTGCTCTGGCTGCAAGCAGGTGCTGACACAGGTGGCTGAAACTGGTTCAACTGTCCAAC  
ACGGATTCCGAGGTAAAGTGACCAACGTGGGTTACGGCTGCGAATGACAGCGGCACT  
ATCTTTGCAACCGACTACGATGGCACCACTGTGACCAACGCGGGCGCAGAGACTGTGACT  
TACAAAGATGCTTCCGGTAAACAGCACCCTGCGGCTGTCACTGGGTGGCTCTGATGGC  
AAAAACCAATCTGGTTACCGCCGCTGACGGCAAAACGTACGGTGCGACTGCACTGAATGGT  
GCTGATCTGTCCGATCTTAATAACACCGTTAAATCTGTTGCAGACAACGCTAAACCGTTG  
GCTGCCCTGGATGATGCAATTGCGATGGTCGACAAATCCGCTCCCTCCGTTGCGGTG  
CAAAACCGTCTGGATTCCGAGTCAACCACTGAAACACCACTACCAACTGTCTGAA  
GCGCAGTCCCGTATTACGAGCGCGACTATGCGACCGAAGTGTCCCAACATGTCGAAAGCG  
CAGATTATCCAGCAGGCAAGTAACTCCGTGCTGTCCAAGCTAACCAAGTTCCGACAGCAG  
GTTCTGTCTCTGCTGACAGGTTAA

Figure 46

69/96

AACAAAAACAGTCTGCGCTGTGCACTTCTATCGAGCGCCTCTTCTGGT  
 CTGCGTATTAACAGCGCTAAAGATGACGCGCGGGCCAGGCGATTGCTAACCGCTTTACT  
 TCTAACATCAAAGGTCTGACTCAGGCCGACGTAAAGCCAAACGACGGTATTCTCTGGCG  
 CAGACCGCTGAAGGCGCGCTGTGAGAGATTAAACAACCTTGCAGCGTATTCTGAACTG  
 ACGGTTGAGGCCTCTACGGCAGCAACTCTGATTCGACCTGTCTTCTATTGAGGACGAA  
 ATCAAATCCCGTCTTGATGAAATTGACCGTGTATCTGGTCAGACCCAGTTCAACGCGTGTG  
 AACGTGCTGTGAAAAACGATTGATGAAGATTCAGATTGGTGCCAATGATAACCCAGAG  
 ATCAGCATTTGGCTTGCAACAAATCGACAGTACCACTTTGAATCTGAAAGGATTTACCGTG  
 TCCGGCATGGCGGATTTACGCGCGGCGAAACTGACGGCTGCTGATGGTACAGCAATTGCT  
 GCTGCGGATGTCAGGATGCTGGGGGTAAACAAGTCAATTTACTGTCTTACACTGACACC  
 GCGTCTAACAGTACTAAATATGCGGTCGTTGATTCTGCAACCGGTAAATACATGGAAGCC  
 ACTGTAGCCATTACCGGTACGGCGCGCGCGGTAACTGTTGGTGACGCGGAAGTGGCGGGA  
 GCCGCTACAGCCGATCCGTTAAAGCACTGGATGCCCAATCGCTAAAGTCGACAAATTC  
 CGCTCTCCCTCGGTGCGGTTCAAAACCGTCTGGATTCTGCGGTACCAACCTGAACAAC  
 ACCACCACCAACCTGTCTGAAGCGCAGTCCCGTATTCAGGACGCCGACTATGCGACGGAA  
 GTGTCCAACATGTGAAAGCGCAGATTATCCAGCAGGCCGGTAACCTCCGTGCTGCCAAA

Figure 47

70/96

ATGGCACAAGTCATTAATAACCAACAGCCTCTCGCTGATCACTC  
 AAAATAATATCAACAAGAACCGAGTCTGCGCTGTCGAGTTCTATCGAGCGTCTGTCTCTG  
 GCTTGCGTATTAAACAGCGCGAAGGATGACGCGCGGTCAGGGCGATTGCTAACCGGTTT  
 CCTCTAACATTAAAGGCTCTGACTCAGGCTGCAAGTAAACGCAACGAGGTATTCTGTTG  
 CACAGACCACTGAAGGCGCGCTGTCCGAAATCAACAACCTTACAGCGTATCCGTGAAC  
 TGACGGTTTCAGGCTTCTACCGGACTAACTCCGATTCCGATCTGGACTCCATTACGGACG  
 AAATCAATCCCGTCTGGACGAAATTGACCGGTATCCGGTCAAAACCGATTCAACGCTG  
 TGAACGTAAGTGGCGAAAGACGGTTCGATGAAAAATTCAGTTGGTGCAATGACGGCCAGA  
 CTATCACGATTGATCTGAAGAAAATTGACTCAGATACGCTGGGCTGAATGGTTTCAACG  
 TTAATGGCAAAGGCATAATTGCGAACAAGCTGTACAGTCAGCGATCTGACCGCTGCTG  
 GTGCAACGGGAACAGGTCTTATGCTGTGACCACAAACAATACAGCACTCAGCGCTAGCG  
 ATGCACTGTCTCGCCTGAAAACCGAGATACAGTTACTACTAGGCTCGAGTGTGCGA  
 TCTATACTTATGATGCGGCTAAAGGGAACCTCACCCTCAAGCAACAGTTGCGAGTGGCG  
 ATGTTGTTAACTTTGCGAATACCTGAAACAGCGGCTGGCACTACTGCATCAGGTGTTT  
 ATACTCGTAGTACTGGTGATGTGAAGTTTGATGTAGATGCTAATGGCGATGTGACCATCG  
 GTGGTAAAGCCCGCTACCTGGACGCCACTGGTAACCTATCTACAAACAACCCCGGCAATG  
 CATCTTCAGCGAAAATTGTCGATCTGTTTGCTAGCGGTAGTACCTTAGCGACAACCTGGTT  
 CTATCCAGCTGTCTGGCACAACCTATAACTTTGGTGCAGCGCGCAACTTCTGGCGTAACCT  
 ACACAAAACCTGTAAGCGCTGATACTGTACTGAGCACAGTGCAGAGTGTGCAACGGCTA  
 ACACAGCAGTTACTGGTGCACAATTAAGTATAATACAGGTATTCACTGCAACGGCGT  
 CCTTCGGTGGTGAATACTAATGGTGTGGTAATTCGAATGACACCTATACATGATGACG  
 ACAAGAGCTCACCAACCCGATCTTACACTATCAACTACAACCTGATAAGGATACCG  
 GTACAGTAAGTGTAGCTTCAAATGGCGCAGGTGCAACTGGTAAATTTGCAGCTACTGTTG  
 GGGCACAGGCTTATGTTAACTCTACAGGCAAACTGACCACTGAAACCAACAGTGCAGGCA  
 CTGCAACCAAGATCTCTGCGTGCCTGGATGAAGCTATCAGCTCCATCGACAACATCC  
 GTTCATCCCTGGGTGCTATCCGAAACCGCTCTGATTCCGCGGTTACCAACCTGAACAACA  
 CCATCTACCAACCTGTCCGAAGCGCAGTCCCGTATTACGAGCCGCACTATGCGACCGAAG  
 TGTCCAACATGTGCAAGGCGCAGATTACCGACAGCGCGTAACTCCGTGCTGCGCAAAAG  
 CCAACCAAGTACCGCAGCAGGTCTGTCTGTCTGTCAGGGTTAA

Figure 48

71/96

ATGGCACAAGTCATTAATACCAACAGCCTCTCGCTGATCAC  
 TCAAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATCGAGCGTCTGTCTTC  
 TGGCTTGCSTATTAAACGCGCGAAGGATGACGCGCGCAGGTCCAGGCGATTGCTAAACCGTTT  
 TACTTCTAATAATTAAGGCGCTGACTCAGGCTGCACGTAAACGCCAATGACGGTATTTCTGT  
 TGCACAGACCACTGAAGGCGCGCTGTCCGAAATCAACAACAATTACAGCGTGTGGGTGA  
 ACTGACCGTTTCAAGGCGACCCGCTACCAACTCCAGTCTGATCTGGACTCTATCCAGGA  
 CGAAATCAAAATCCCGTCTGGACGAATGACCGGTATCCCGTCTGAGACTCAGTTCAACGG  
 CGTGAACGTACTGGCAAAAGACGGTTCATGAAAATTCAGGTTGGCGCGAATGATGGCCA  
 GACCATCACTATCGACCTGAAGAAGATTGACTCTTCTACGTTGAAACTGACTGGTTTTAA  
 CGTGAATGGTTCTGGTTCTGTGGCGAATACTGCGGCGACTAAAGACGAATGGCTGCTGC  
 TGCTGCGGCGGCGGGTACAACCTCCTGCTGTGGTACTGACGGCGTGACCAAAATATACCGT  
 AGACGCGAGGGCTTAACAAAGCCACAGCAGCAAAACGTTGTTGCAAACTTGCAGATGGTGC  
 TGTGTTGATGCTAGCATTTCCAACGGTTTTGGTGACGACGACCAAGACTACACCTA  
 CAATAAAGCTACAATAATGATTCACTTTCAATGCCAGCATTTGCTGCTGGTGTGCGGCGG  
 TGATAGTAACAGCGCAGCTCTGCAATCCTTCTGACTCCAAAAGCAGGTGATACAGCTAA  
 CCTGAGCGTCAAAATCGGTACGACATCTGTTAATGTTGTTCTGGCGAGCGATGGCAAAAT  
 TACAGCGAAAGATGGCTCAGCTCTGTATATCGACTCAACGGGTAACTGACTCAGAACAG  
 CGCAGGCACTGTAAACAGCAGCAACCCCTGGATGGACTGACCAAAACCATGATGCGACAGG  
 AGCTGTTGGTGTGATATCAGCAGCGCAGATGGCGCAACTATCTCTGCGCAGGCTCTGC  
 TAACGCGCAACAGGTACTCAATCAGGTGCAATTAACCTGAAAATGTTGCTATCAGTGC  
 TGATGCTCTGCAGTCTGCTCGAAAGGTACTGTTATCAATGTTGATAATGGTGTCTGATGA  
 TATTTCTGTTAGTAAACCGGGTGTCTGTTACTACCGAGGTGCGCCTACTTATCTGATG  
 CTGATGGTAAATTAACGACAACCAACCGTTGATTAATTTCTGCAAACTGATGGCAGCG  
 TAACCAATGGTTCTGGTAAAGGGTTTACACCGATGCGAGTGGTAAATTAACCTACCGACG  
 CTGCAACCAAGCGCGCAACCAACCGGATCCGCTGAAAGCCCTTGATGACGCAATCAGCC  
 AGATCGATAAGTTCCGTTTCACTCCCTGGGTGCTATCCAGAACCCTGTGGATTCGCGGTTA  
 CCAACCTGAACAACCACTACCAACCTGTCCGAAGCGCAGTCCCGTATTCAGGACCGCG  
 ACTATGCGACCGAAGTGTCCAATATGTGAAAGCGCAGATCATCCAGCAGGCGCGTAACT  
 CCGTGTGGCAAAAGCTAACCGAGTACCGCAGCAGGTCTCTCTCTGCTGAGGGTTAA

Figure 49

72/96

AACAAATCTCAGTCTTCTCTTAGCTCTGCTATTGAGCGTCTGCTCTTCTGGT  
CTGCGTATTAAACAGCGCAAAAGACGATGCAGCAGGTCAGGCGATTGCTAACCGTTTTACG  
GCAAAATATTAAGGTCTGACCCAGGCTTCCCGTAACGCGAATGATGGTATTTCTGTTGCG  
CAGACCACTGAAGGTGCGCTGAATGAAATTAACAACAACCTGCAGCGTATTCGTGAACTT  
TCTGTTTCAGGCAACTAACGGTACTAACTCTGACAGCGATCTTTCTTCTATCCAGGCTGAA  
ATTACTCAACGTCTGGAAGAAATTGACCGTGTATCTGAGCAAACTCAGTTTAACGGCGGTG  
AAAGTCCTTGCTGAAAAATAATGAAATGAAATTCAGGTTGGTGCTAATGATGGTGAAACC  
ATCACTATCAATCTGGCAAAAATTGATGCGAAAACTCTCGGCCTGGACGTTTTATATC  
GATGGCGCGCAGAAAGCAACCGGCAGTGACCTGATTTCTAAATTTAAAGCGCAGGTA  
GATAATTATCAAATTAACGGTACTGATACTATCTGTTAATGTAGATAGTGGAGTAGTA  
CAGGATRAAGATGGCAACAAGTTTATGTGAGTGCTGCGGATGGTTCACTTACGACCCAGC  
AGTGATACTCAATTCAAGATTGATGCAACTAAGCTTGCGAGTGGCTGCTAAAGATTAGCT  
CAAGGTAATAAGATTGTCTACGAAGGTATCGAATTTACAAATACCGGCACCTGGCGCTATA  
CCTGCCACAGGTAATGGTGAATTAACCCCAATGTTGATGGTAAGGCTGTTGAATTCAC  
ATTTCCGGGAGTGCTGATACATCAGGTACTAGTGCAACCGTTGCCCTACGACAGCCCTA  
TACAAAAATAGTGACGGCAATTGACTGCAACAAAAGTTGAAAAATAAGCAGCGACACTA  
TCTGATCTTGATCTGAACGCTGCCAAGAAAAAGCAAGCAGTCTAGTTGTTAACGGTGCA  
ACTTACGATGTTAGTGACGATGGTAAACGATAACGGAGACTGCTTCTGGTAACAATAAA  
GTCATGTATCTGAGCAAAATCAGAAAGGTGGTAGCCCGATTCTGGTAAACGAAGATGCAGCA  
AAATCGTTGCAATCTACCACCAACCCGCTCGAAACTATCGACAAGCATTGGCTAAAGTT  
GACAATCTGCGTTCTGACCTCGGTGCAGTACAAAACCGTTTCGACTCTGCCATCACCAC  
CTTGGCAACCCGTAAACAACCTGTCTTCTGCGCGTAGCCGTATCGAAGATGCTGACTAC  
GCGACCGAAGTGTCTAACATGTCTCGTGCGCAGATCTGCAACAAGCGGGTACCTCTGTT CTGGCACAG

Figure 50

73/96

ATGGCACAAGTCATTAATACCAACAGCCTCTCGCTGATCAC  
 TCAAAATAATATCAACAAGAACCAAGTCTGCGCTGTCGAGTTCTATCGAGCGTCTGTCTTC  
 TGGCTTGGTATTAACAGCGCGAAGGATGACGAGCGGGTCAGGCGATTGCTAACCGTTT  
 CACCTCTAACATTAAGGCGTGACTCAGGCGGCCCGTAACGCCAACGACGCTATCTCCGT  
 TGCGCAGACCAACGAAGGCGCGCTGTCCGAAATCAACAACAACTTACAGCGTGTGCGTGA  
 ACTGACGGTACAGGCCACTACCGGTACTAACTCTGAGTCTGATCTGTCTTCTATCCAGGA  
 CGAAATTAATCCCGTCTGGATGAAATTGACCGCGTATCTGGTCAGACCCAGTTCAACGG  
 CGTGAACGTGCTGGCAAAAAATGGCTCCATGAAAAATCCAGGTTGGCGCAATGATAACCA  
 GACTATCACTATCGATCTGAAGCAGATTGATGCTAAAACTCTTGGCCTTGATGTTTTAG  
 CGTTAAAAATAACGATACAGTTACCACTAGTGCTCCAGTAACCTGTTTTGGTGCTACCAC  
 CACAACAATATTAACCTTACTGGAATTACCCCTTCTACGGAAGCAGCCACTGATACTGG  
 CGGAACATAACCCAGCTTCAATTGAGGGTGTTTATACTGATAATGGTAATGATTACTATGC  
 GAAAAACACCGTGGTGATAACGATGGGAAGTATTACGCAGTAACAGTTGCTAATGATGG  
 TACAGTGACAATGGCGACTGGAGCAACGGCAATGCAACTGTGAATCTGATGCAAACTACTAC  
 TAAAGCTACAACCTATCACTTCAGGCGGTACACCTGTTTCAAGATTGATAAATCTGCAGGTTT  
 CGCAACTGCCAACCTTGGTGCTGTTAGCTTAGTAAAACTGCAGGATTCGAAGGGTAATGA  
 TACCGATACATATGCGCTTAAAGATACAAATGGCAATCTTTACGCTCGGATGTGAATGA  
 AACTACTGTTGCTGTTTCTGTTAAACTATTACCTATACTGACTCTTCGSGTGCCGCCAG  
 TTCTCCAAACGCGGTCAAACCTGGGCGGAGATGATGGCAAAACAGAAGTGGTGCATATTGA  
 TGTGAAAAACATACGATTCTGCCGATTAAATGGCGGTAATCTGCAAAACAGGTTTGACTGC  
 TGGTGGTGAGGCTCTGACTGCTGTTGCAAAATGGTAAACCACCGGATCCGCTGAAAACGCT  
 GGACGATGCTATCGCATCTGTAGACAAATCCGTTCTTCCCTCGGTGCGGTGCAAAACCG  
 TCTGGATTCCGCGGTTACCAACCTGAACAACCACTACCAACCTGTGTGAAGCGCAGTTC  
 CCGTATTGAGGACGCCGACTATGCGACCGAAGTGTCCTAATATGTCGAAAAGCGCAGATCAT  
 CCGCAGCGCCGGTAACTCCGTGTTGGCAAAAGCTAACAGGTAACGCGCAGGTTCTGTCT  
 TCTGCTGCAGGGTTAA

Figure 51

74/96

ATGGCACAAAGTCATTAATACCAACAGCCTCTCGCTGATCACT  
 CAAAAATAATATCAACAAGAACCAGTCTCGCGTGTGAGTTCATATCGAGCGTCTGTCTTCT  
 GGCTTGCCTATTAAACAGCGCGAAGGATGACGCCGAGGTGAGCGGATTCGTAAACCGTTTT  
 ACTTCTTAACATTAAAGGCCGTGACTCAGGCTGCACGTAAACGCCAACGACGGTATTTCTGTT  
 GCGCAGACCCGGAAGGCGCGTGTCTGAAATCAACAACAACTTACAGCGTATTCTGTGAA  
 CTGACGGTTCAGGCTTCTACCGGGACTAACTCTGATTCGGATCTGGACTCCATTTCAGGAC  
 GAAATCAAAATCCCGTCTGGACGAAATTGACCGCGTATCCGGTCAAAACCAAGTTCACCGT  
 GTGRACGTACTGGCGAAAGACGGTTCGATGAAATTCAGGTTGGTGCGAATGACGGCCAG  
 ACTATCACTATTGATCTGAAGAAAAATGACTCTGATACGCTGGGGCTGAATGGTTTTAAC  
 GTTAACGGCAAGGTACTATTGCGAACAAAGCGGCAACCTATTAGTATCTGGCGGCGAGC  
 GGGCGGAATGTTACTAACTCAAGCAATATTGTTGTACGACAAAGTTCATGCCTTGGAT  
 GCAGCGACTGCATTAGCAAACTCAAAGATGGTGATTCTGTTGCCGTTGCTGCTCAGAAA  
 TATACTTTATAACGCATCGACCAATGATTTTACGACAGAAAATACAGTAGCGACAGGCACT  
 GCAACGCACAGATCTTGGCGCTACTCTGAAGGCTGCTGTGGGCAGAGTCAATCAGGTACA  
 TATACCTTTGCAAAATGGTAAAGTTAACTTTGATGTTGATGCAAGCGGTAATATCACTATT  
 GGGCGGAAAAAGGCTTTCTTGGTTGGTGGAGCGCTGACTACTAACGATCCACCGGCTCC  
 ACTCCAGCAACGATGTCTTCCCTGTTTAAAGGCCGCGGATGACAAAGATGCCGCTCAATCC  
 TCGATTGATTTTGGCGGAAAAAATACGAATTTGCTGGTGGCAATTCTACTAATGGTGGC  
 GGCGTTAAATCAAAGACACGGTGTCTTCTGACGCGCTTTTGGCTCAGGTTAAAGCGGAT  
 AGTACTGCTAATAATGTAATAATCACCTTTAACAATGGTCTCTGTCACTTCACTGCATCG  
 TTCCAAAATGGTGATCTGGCTCCGCGGCATCGAATGACGCGCTACATTGATAGCGAAGGC  
 GAACTGACAACTACTGAATCCTACAACACAAATATTCCGTAGACAAAGACACGGGGGCT  
 GTAAGTGTTACAGGGGGGAGCGGTACGGGTAATACGCCGCAACGTGGGTGCTCAGGCT  
 TATGTAGGTGCAGATGGTAAATTAACCACGAATACTACTAGTACCGGCTCTGCAACCAAAA  
 GATCCACTAAATGCGCTGGATGAGGCAATTGCATCCATGACAAATTCGGTCTCTCCCTG  
 GGGGCTATCCAGAACCGTCTGGATTCCGCAGTCACCAACCTGAACAACACCACCTACCAAC  
 CTGTCTGAAGCGCAGTCCCGTATTGAGGACGCCGACTATGCGACCGAAGTGTCCAACATG  
 TCGAAAGCGCAGATCATCTCAGCAGGCCGGTAACTCCGTGTTGGCAAAAGCTAACCAAGTA  
 CCGCAGCAGGTTCTGTCTCTGCTGCAGGGTTAA

Figure 52



75/96

AACAGAACCAGTCTGCGCTGTCGAGTCTATCGAGCGTCTGTCT  
 TTCTGGCTTGCGTATTAAACAGCGCGAAGGATGACGCCGCGGTGAGGCGATTGCTAACCG  
 TTTTACTTCTAACATTAAAGGCCTGACTCAGGCTGCACGTAACGCCAACGACGGTATTTC  
 TGTTCGCGAGACCACCGAAGCGCGCTGTCCGAAATTAACAACACTTACAGCGTGTGCG  
 TGAGCTGACTGTTTCCAGGCGACCCCGTACTAACTCTGAGTCTGACCTGTCTTCTATCCA  
 GGACGAAATCAAATCTCGCCTGGAAGAGATTGATCGTGTTTCAAGTCAAGTCAATTTAA  
 CGGCGTGAATGTTTGGCTAAAGATGGGAAAAATGAACATTACAGTTGGGGCAAGTGATGG  
 ACAGACTATCACTATTGATCTGAAAAAGATCGATTCACTTACACTAAACCTCTCCAGTTT  
 TGATGCTACAAACTTGGGCACCACTGTTAAAGATGGGGCCACCATCAATAAGCAAGTGGC  
 AGTAGATGCTGGCGACTTTAAAGATAAAGCTTCAGGATCGTTAGGTACCCATAAATTAGT  
 TGAGAAAGACGGTAACTACTATGTAATGACACTAAAAGTAGTAAGTACTACGATGCCGA  
 AGTAGATACTAGTAAGGTGAAATTAACCTTCAACTCTACAAATGAAAGTGAAGTACTCC  
 TACTGCAGCGACGGAAGTAACTACTGTTGGCCGCGATGTAAAAATTGGATGCTTCTGCACT  
 TAAAGCCAACCAATCGCTTGTCGTGTATAAAGATAAAGCGGCAATGATGCTTATATCAT  
 TCAGACCAAGATGTAAACACTAATCAATCAACTTCAATGCCGCTAATATCAGTGATGC  
 TGGTGTTTATCTATTGGTGCACTCTACAACCGCGCCAAGCAATTTAACAGCTGACCCGCT  
 TAAGGCTCTTGATGATGCAATTGCATCTGTGTATAAATTCGCTCTTCTCGGTGCCGT  
 TCAGAACCGTCTGGATTCTGCCATTGCCAACCTGAACAACCACTACCAACCTGTCTGA  
 AGCGCAGTCCCGTATTCAGGACGCTGACTATGCGACCGAAGTGTCCAACATGTCCGAAAGC  
 GCAGATTATCCAGCAGCCGGTAACTCCGTGCTGGCAAAA

Figure 53

76/96

ATGGCACAGTCATTAAATACCAACAGCCTCTCGCTGATCACTCAAAA  
 TAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATCGAGCGTCTGTCTTCGGCTT  
 CGGTATTAAACAGCGCGAAGGATGACGACGCGGTGAGCGGATTGCTAACCGTTTCACTTC  
 TAACATTAAAGGCCCTGACTCAGGCTGCACGTAACGCTAACGATGGTATCTCTCTGGCGCA  
 GACCACTGAAGCGCACTGTCTGAGATTAAACAACAATTACACGTTGTCGGTGAGTTGAC  
 TGTACAGCGCACACCGGTACTAACTCTGATTCTGACCTGGCTTCTATTCAAGACGAAT  
 CAAATCCCGTTTGTCTGAAATTGACCGGTATCCGGCGAGACCCAGTTCAACGGCGTGAA  
 CGTATTGCTAAAGATGGCTCCCTGAAAAATTCAAGTTGGCGCAATGATGGTCAGACTAT  
 CTCTATCGACCTGAAGAAAATTGACTCTGATACTCTGGGTTTGAATGGTTTCAACGTTAA  
 TGGTTCTGGTACCATTGCAACAAAGCGGCCACAATCAGTGACTTGACTGCTCAGAAAGC  
 CGTTGCAACGGTAATGGTACTTATAAAGTTACAAC TAGCAACGCTGCACTTACTGCATC  
 TCAGGCATTAAAGTGAAGCTGAGTGATGGCGTACTGTAGATATTGCAACCTATGCTGGTG  
 TACAAGTTCAACAGTTAGTTATAAATACGACGAGATGCAGGTAACTTCAGTTATAACAA  
 TACTGCAACAAAACAAGTGCTGCGGTGGAACCTCTGGCAGATACTCTTCTCCCGGCAGC  
 TGGCCAGACTAAAACCGGTACTTACAAGGCTGCTACTGGTGATGTTAACTTTAATGGTGA  
 CGCAACTGGTAATCTGACAATTGGCGGACAGCAAGCCTACCTGACTACTGATGGTAACCT  
 TACAACAAACAACCTCCGGTGGTGCAGCTACTGCAACTCTTAAAGAGCTGTTTACTCTTGC  
 TGGCGATGGTAAATCTCTGGGAACGGCGGTACTGCTACCGTTACTCTGGATAATACTAC  
 GTATAATTTCAAAGCTGCTGCGAACGTTACTGATGGTGCTGGTGTCATCGCTGCTGG  
 TGTAACTTATACAGCACTGTTTCTAAAGATGTCATTCTGGCAACAAGTCAATCTGCAAG  
 TCAGGCAGCAGCAACCGCTACCGACGGTGATCTGTGCAACGATCAACTATAAATCTGG  
 TGTCATGATCGGTTCCGCTACCTTTACCAATGGTAAAGGTACTGCGGATGGTATGACTTC  
 TGGTACAACCTCCAGTCGTAGCTACAGGTGCTAAAGCTGTATATGTTGATGGCAACAATGA  
 ACTGACTTCCTGCTGATCTTACGATACGACTTACTCTGTCAACGCAGATACAGGCGCAGT  
 AAAAGTGGTATCAGGTACTGGTACTGGTAAATTTGAAGCTGTTGCTGGTGGGATGCTTA  
 TGTAAGCAAAGATGGCAAAATTAAACGACAGAAACCAAGTGCAGGCACTGCAACCAAGA  
 TCCTTTGGCTGCCCTGGATGCTGCTATCAGCTCCATCGACAAATTCGGTTCCTCCCTGGG  
 TGCTATCCAGAACCGTCTGGATTCCGAGTCACCAACCTGAAACAACCACTACTAACT  
 GTCTGAAGCGCAGTCCCGTATTACGACGCCGACTATGCGACCGAAGTGTCCAATATGTC  
 GAAAGCGCAGATCATCCAGAGGCCGGTAACCTCTGTGTTGGCAAAAGCTAACAGGTACC  
 GCAGCAGGTTCTGTCTCTGCTGCAGGGTTAA

Figure 54

77/96

ATGGCACAAGTCATTAAATACCAACAGCC  
 TCTCGCTGATCACTCAAATAATATCAACAAGAACCAGCTGCGCTGTCGAGTTCTATCG  
 AGCGTCTGTCTTCTGCGCTTGCCTATTAAACAGCGCGAAGGATGACGCCGAGGTGAGGCGA  
 TTGCTAACCGTTTTACTTCTAACATTAAAGGCCCTGACTCAGGCTGCACGTAACGCCAACG  
 ACGGTATTTCGTGTGCACAGACCACTGAAGGCCGCTGTCCGAAATCAACAACAATTAC  
 AGCGTATTCTGTAACCTGACGGTTTACGGCTTCTACCGGGACTAACTCTGATTTCGGATCTGG  
 ACTCCATTGAGSACGAAATCAAATCCCGTCTCGACGAAATTGACCGGTTTTCCGGTCAGA  
 CCCAGTTCAACGGCGTGAACGTGCTGGCGAAAGACGGTTTCGATGAAGATTTCAGGTTCGCG  
 CGAATGACGGCGAGACCATCTCTATCGATTTCAGAAAAATTGATTCTTCAACGCTGGGAT  
 TGAAGGTTTTCTCGGTATCAGGGAACGCATTAAAGTTAGCGATGCGATAACTACAGTTC  
 CTGGTGCTAATGCTGGCGATGCCCGGTTACGGTTAAATTTGGTGCGAACGATACCGCTG  
 CTGCCGCAATGGCTAAACATTGGGAATAAGTGATACATCAGGCTTGTCCCTACATAACG  
 TACAAAGCGCGGATGGTAAAGCGACAGGAACCTATGTTGTTCAATCTGGTAATGACTTCT  
 ATTCGGCTTCGGTTAATGCTGGTGGCGTTGTTACGCTTAATACCACCAATGTTACTTTCA  
 CTGATCTTCGGAACGGTGTACCAAGCAACACAGACAGGTGACGCTATCAAGGTCACGA  
 CGAATAGTGCTGGCGCGGCTGTTGGCTATGTTACTATTCAAGGCAAGATTACCTTGCTG  
 GTGCAGACGGTAAGGATGCAATTGAAAACGGTGGTGACGCTGCAACAAATGAAGACACAA  
 AAATCCAACTTACCGATGAACTCGATGTTGATGTTCTGTAAAAACAGCGGCAACAGCAA  
 CATTTCCTGGTACTGCAACCAACGATCCGCTGGCACTTTAGACAAAGCTATCTCGCAAG  
 TTGATACTTTCGCTCCTCCCTCGGTGCCGTACAAAACCGCTCTGGATTCTCGGTCACCA  
 ACCTGAATAACACCACCACCAACCTGTCTGAAGCGCAGTCCCGTATTGAGGACGCGGACT  
 ATGCGACCGAAGTGTCCAACATGTGAAAAGCGCAGATCATCCAGCAGGCGGGTAACTCTG  
 TGCTGTCTAAAGCTAACCCAGGTACCGCAGCAGGTTCTGTCTCTGCTGCAGGTTAA

Figure 55

78/96

CTTCTCTTAGCTCTGCTATTGAGCGTCTGTCTTCTGTCTGCTGATTAAACAGCGCAAAAG  
 ACGATGCGACAGGTGAGCGGATGCTAACCGTTTACGGCAAAATATAAGAGTCTGACCC  
 AGGCTTCCCGTAACGCGAATGATGGTATTTCTGTTGCGCAGACCACTGAAGGTGCGCTGA  
 ATGAAATTAAACAACACCTGACGCGTATTCGTGAACTTCTGTTCAGGCAACTAACGGTA  
 CTAACCTGACAGCGATCTTTCTTCTATCCAGGCTGAAATTACTCAACGCTCTGGAAGAAA  
 TTGACCGGTGATCTGAGCAAACTCAGTTTAAACGGCTGAAAGTCTTGTGTAATAATG  
 AAATGAAATTACAGTTGGTGCTAATGATGGTGAACCACTTACCTGCCGCCACGATTAG  
 ATACAACACTCAGTTAGTAACGTGGAATCTTCAATTCAGAAATGACCTTTCTCCAGCC  
 CGCTGCAAAATTCAGACGGTGTCTGATAATTCAGCGTGGAGTGGCGGCGGCATTCTGTATA  
 ATCCTGCCGCCAGTCATTAAATTTTCTGGCATGAACGATATCGCTGAACCACTGCTC  
 ATTCAAACTTATCGCGAAATCGTCGGTAAAGCTCTCAATAAATCCGTTCTGCGTTGG  
 CTGCGCCGCTGGATTAAAGCGCACTCAACACCATGCTCAAGAGGCCATTGATCCAGTGC  
 ATGCGCAAGTGAACCTCGGCCCTGGTCAGTTCTTATCGTCGGGATAGCTCGCAACAG  
 TGCAATGCTGTCGGAATACGCTGACCTGAACGCTGAAATCCCAAGGCAACAGTGAC  
 CGTCAGGCATTCCTTTGTGAAATCATCGACGCGAGTGAACACTTGATCTCGCACCGGT  
 GGAAAGTGCCTCCATGACGAATCCATCGACCGGTGAGTTGGCGCGCGCGGACGGAG  
 CAGCGCGGACGCTTCTGTTGCGCAGCCCTTACGACGCTCTCTGCGTTTACGCCCGAGGCC  
 ACTGAGGTGATAAAGCCGGTACACGCGCTTATGATTAACTGAAGCCCTTCAGCGCGAG  
 CAACTGCCAAATACGACGGTAGCCAAAACGCTGCGCTCCAGTGCCAGCTCAGTGATCG  
 CCTGATAAATGCGCATCAGCAGCGCGGACGGTGAACCTCATAGCGCGAGTGCACAGGA  
 NNAACCTGTAAGCCTGACGGCACGACGTTGCGACAGACCGGTGCGCATCACACATCAACAT  
 CAGCGCTTCCCGCTTCTGGTCTGTGCTGCTGACTTTCGCCCAAGAGCCACCTGAAGCGCC  
 TCTTTATCCAGCATGGCTTCGGCAAGCAGCTTCTTGAGTCTGGTGTCTCTTCTCAAGC  
 GACTTCAGCGCTTAACCTCAGGCACCTCCATACCGCCATCTTCTTACGCCAGGTGTA  
 AACGTGGCATCGAAATGGCATGCTTCCGCGCAGGTTACGCGCGGTTACCCAGCTTCG  
 GCTTCGCGGAGATACTGATGATCTGTTCTGTCGGAAAAACGCTTCTTCTATGGGATGTCC  
 TCAATGCGCTTATGAAGACATTACTAATACGCGGTGACTAATCAACGGGAGCAGGTG  
 ACCATCACTATCAATCTGGCAAAAATTGATGCGAAAACTTCGCGCTGGACGCTTTAAT  
 ATCGATGGCGCGCAGAAAGCAACCGGAGTGACCTGATTTCTAAATTTAAAGCGACAGGT  
 ACTGATAATTATCAAAATTAACGCTACTGATAACTATACTGTTAATGTAGATAGTGGAGTA  
 GTACAGGATAAAGATGGCAAAACAAGTTTATGTAGTGTGCGGATGGTTCACCTTACGACC  
 AGCAGTGATACTCAATTCAAGATTGATGCAACTAAGCTTGACGTGGCTGCTAAAGATTTA  
 GCTCAAGGTAATAAGATTGTCTACGAAGGTATCGAATTTACAAATACCGGCACTGGCGCT  
 ATACCTGCCACAGGTAAATGGTAAATTAACCGCCAATGTTGATGGTAAGGCTGTGTAATTC  
 ACTATTTGGGGAGTGTGATACATCAGGTACTAGTGCAACCGTTGCCCTACGACAGCC  
 CTATACAAAAATAGTGCAGGGCAATGACTGCAACAAAAGTTGAAAAATGAAGCAGCGACA  
 CTATCTGATCTTGATCTGAACGCTGCCAAGAAAAAGCAAGCAGCTTAGTTGTTAAACGGT  
 GCAACTTACGATGTTAGTGAGATGGTAAACGATAACGGAGACTGCTTCTGGTAACAAT  
 AAAGTCATGTATCTAGCAAAATCAGAAGGTGGTAGCCGATTTCTGGTAACCAAGATGCA  
 GCAAAATCGTTGCAATCTACCAACCAACCGCTCGAAACTATCGACAAAGCATTTGGCTAAA  
 GTTGACAATCTGCGTTCTGACCTCGGTGCGATACAAAACCGTTTCGACTTGCCTACAC  
 AACCTTGGCAACCCGTAACCAACCTGTCTTCTGCCGCTATCGAGCTATCGAAGATGCTGAC  
 TACGCGACCGAAGTGTCTAATATGCTCTGTCGCGAGATCTCGCAACAGCGGTTACCTCT  
 GTTCTGGCACAGCTAAAC

Figure 56

79/96

AACAAAAACCAGTCTGCGCTGTCGACTTCTATCGAGCGCCTCTCT  
 TCTGGTCTGCGCATTAAACAGCGCTAAAGATGACGCTGCGGGCCAGGCGATTGCTAACCGC  
 TTCACTTCTAACATCAAAGGTCTGACTCAGGCGGCACGTAACGCCAACGACGGTATCTCT  
 CTGGCGCAGACCACTGAAGGCGCACTGTCTGAAATCAACAACAACCTTGACGCGTGTTCGT  
 GAAGTACCGGTTTCCAGGCCACTACCGGTACTAACTCTGATTCTGACCTGTCTTCAATCCAG  
 GACGAAATCAAATCCCGTCTCGATGAAATTGACCGCGTATCCGGTCAGACTCAGTTCAAAC  
 GCGGTGAACGTACTGGCAAAAGATGGCTCGATGAAATTCAGGTCCGGTGCAAAATGATGGT  
 CAGACAATCAGCATTGATTTGCAGAAGATTGATTCTTCTACTTTAGGGTTAAATGGTTTTT  
 TCTGTTTCCAAAAATGCAGTATCTGTTGGTGATGCTATTACTCAATTGCGTGGCGAGAGC  
 GCAGCCGATGCACCACTAACCATCAAGTTTGATGATTGATGAAAACTGATTTAAACTG  
 ACCGATGCTTCAGGGTTAAGTCTGCATAAACCCTCAAGATGAAATGGTAATTTAACTAAC  
 CAGTATGTTGTACAGAATGGCGGAAAACTCTACGCTGCTACAGTCGCTGCCAATGTAAT  
 GTTACGCTGAACAAAGCAAATGTAACCTACAGCGATGTGCAAAAGGTTATTGATACCGCA  
 ACGCAGTCAGGCCAGTTAGTTTCAGGTTGGTGACAGATTCTACCGGTACGCCAAAAGCATT  
 GTGCTGTCTCAAGGTAAAAGCTTTGGCATTGATGACGCCGCCCTTGAAGATAAACACTGGT  
 GATGCTACCGCTACTCCACCGGGAACATCTGGGACAACAGTTGTGCGAGCGCTCAATTCTAT  
 CTGAGTACGGGCAAAACTCTGTAGACGCTGATGTAACGGCTTCCACTGAAATTCACAGGT  
 GCTTCAACCAACGATCCACTGACTCTGCTGGACAAGCTATCGCATCTGTTGATAAAATTC  
 CGTTCTCTTTGGGGCGGTACAGAACCGTCTGAGCTCCGCTGTAACCAACCTGAACAAC  
 ACCACCACCACTGTCTGAAGCGCAGTCCCGTATTCAGGACGCCGACTATGCGACCGAA  
 GTGTCCAACATGTCGAAAGCGCAGATTATCCAGCAGGCGAGTAACTCCGTGCTGTCCAAA

Figure 57

80/96

AACAAAAACCACTCTGCGCTGTCGACTTCTATCGAACGCTCTCTTCTGG  
 CCTGCGTATTAACAGTGCAGAAAGATGACGCTGCCGGTCAGGCGATAGCTAACCGTTTCAC  
 CTCTAACATTAAAGGCTGACTCAGGCTGCGCGTAAACGCCAACGACGGTATTTCTCTGGC  
 GCAGACCACAGAAGGTGCGTTGCTGAAATCAACAACAACCTTGCAACGTGTGCGCTGAGTT  
 GACCGTTCAGGCGACGACCGGTACTAACTCTGATTCTGACCTGTCTATCTATTTCAGGACGA  
 AATCAAAATCCCGTCTGGATGAGATTGACCGTGTTCGGTTCAGACCCAGTTCAACGGCGT  
 GAATGTACTGGCAAAAGACGGTTCGATGAAGATTGAGTTGGCGCGAATGATGGCCAGAC  
 TATTAGCATTGATTACAGAAAAATTGACTCTTCTACATTAGGGTTGAATGGTTTCTCGGT  
 TTCTGCTCAATCACTTAAAGTTGGTGATTCAATTACTCAAAATACAGGAGCCGCTGGGAC  
 AAAACCTGTTGGTGTGATTTCAGTCTGTTGCGAAAGATCTGACTACTGCGACAGGTAA  
 AACTGTGCGATGTTCCAGCCTGACGTTACACAACCCCTGGATGCGAAAGGGGCTGCCAC  
 CGCACAGTTCGTCGTTCAATCCGGTAGTGATTCTACTCCGCGTCCATTGACCATTGCAAG  
 TGGTGAAGTGACGTTGAATAAAGCCGATGTCGAATACAAAGACACCGATAATGGACTAAC  
 GACTGCGAGTACTCAGAAAGATCAGCTGATTAAAGTTGCCGCTGACTCTGACGGCGCGGC  
 TCGGGGATATGTAACATTCAGGGTAAAAACTACGCTACAACGGCTCCAGCGCGCTTAA  
 TGATGACACTACGGCAACAGCCACAGCGGAACAAGTTGTTGTTGAATTATCTACAGCAAC  
 TCGACTGCGCAGTTCTCAGGGGCTTCTTCTGCTGATCCACTGGCACTTTTAGACAAAGC  
 CATTGCACAGGTTGATACTTTCCGCTCCTCCCTCGGTGCCGTTCAAAACCGCTGGAATC  
 TGCGGTAAACCAACCTGAACAACACCACCAACCTGTCTGAAGCGCAGTCCCGTATTCA  
 GGACGCCGACTATGCGACCGAAGTGTCTAACATGTGCAAAAGCGCAGATCATCCAGAGGC  
 GGGTAACTCTGTCTGTCTAAA

Figure 58

81/96

ATGGCACAAAG TCATTAATAC CAACAGCCTC TCGCTGATCA CTCAAAATAA TATCAACAAG  
 AACCACTCTG CGCTGTCGAG TTCATCAGAG CGTCTGTCTT CTGGCTTGCG TATTAAACAGC  
 GCGAAGGATG ACGCCGCGGG TCAGGCGATT GCTAACCGTT TTACTTCTAA CATTAAAGGC  
 CTGACTCAGG CTCACGCTAA CGCCAACGAC GGTATTCTCT TTGCACAGAC CACTGAAGGC  
 GCGCTGTCCG AAATCAACAA CAACCTTACAG CGTATCCGTG AGCTGACGGT TCAGGCTTCT  
 ACCGGGACTA ACTCTGATTG GGATCTGGAC TCCATTGAGC ACGAAATCAA ATCCCGTCTC  
 GACGAAATTG ACCGCGTATC CGCTCAGACC CAGTTCAAAG GCGTGAACGT ACTGGCAAAA  
 ACGGGTTCGA TGAATAATCA GGTGCGTGGC AATGACGGTG AAATATCAC TATCGACCTG  
 AAGAAAATCG ATTCTGATAC TCTGGGCTCG AATGGTTTAA ACGTAAATGG TAAAGGTACT  
 ATTACCAACA AAGCTGCAAC GGTAAGTGAT TTAACCTCTG CTGGCGCGAA GTTAAACAC  
 CACGACAGGT CTTTATGATC TGAATAACCGA AAATACCTTG TTAACCTACG ATGCTGCATT  
 CGATAAATTA GGGAAATGGC ATAAAGTCAC CGTTGGCGGC GTAGATTATA CTTACAACGC  
 TAAATCTGGT GATTTTACTA CCACCAAAAT TACTGCTGGT ACGGGTGTAG ACGCCGCGGC  
 GCAGGCTACT GATTGAGCTA AAAAAAGTGA TCGTTAGCT GCCACCTTTC ATGCTGATGT  
 GGGTAAATCT GTTAATGGTT CTTACACCAC AAAAGATGGT ACTGTTTCTT TCGAAACCGA  
 TTCAGCAGGT AATATCACCA TCGGTGGAAG CCAGGCATAC GTAGACGATG CAGGCAACTT  
 GACGACTAAC AACGCTGGTA GCGCAGCTAA AGCTGATATG AAAGCGCTGC TTAAGCGCGC  
 GAGCGAAGGT AGTGACGGTG CCTCTCTGAC ATTCAATGGC ACTGAATATA CTATCGCAAA  
 AGCAACTCCT GCGACAACCT CTCAGTAGC TCCGTTAATC CCTGGTGGGA TTACTTTATCA  
 GGCTACAGTG AGTAAAGATG TAGTATTGAG CGAAACCAAA GCGGCTGCCG CGACATCTTC  
 AATTACCTTT AATTCGGGTG TACTGAGCAA AACTATTGGG TTTACCGCGG GTGAATCCAG  
 TGATGCTGCG AAGTCTTATG TGGATGATAA AGGTGGTATT ACTAACGTTG CCGACTATAC  
 AGTCTCTTAC AGCGTTAACA AGGATAACGG CTCTGTGACT GTTGCGGGGT ATGCTTCAGC  
 GACTGATACC AATAAAGATT ATGCTCCAGC AATTGGTACT GCTGTAAATG TGAACCTCCG  
 GGGTAAATC ACTACTGAGA CTACCAAGTG TGGTTCTGCA ACGAACCAAC CGCTTGCTGC  
 CCTGGACGAC GCTATCAGCT CCATCGACAA ATTCGCTTCT TCCCTGGGTG CTATCCAGAA  
 CCGTCTGGAT TCCGAGTCA CCAACCTGAA CAACACCATT ACCAACCTGT CTGAAGCGCA  
 GTCCCGTATT CAGGACGCCG ACTATGCGAC CGAAGTGTCC AACATGTCGA AAGCGCAGAT  
 TATCCAGCAG GCGGTTAACT CCGTGTCTGC AAAAGCCAAC CAGGTACCGC AGCAGGTTCT  
 GTCTCTGCTG CAGGGTTTAA

Figure 59

82/96

ATGGCACAAG TCATTAATAC CAACAGCCTC TCGTGATCA CTCAAAATAA TATCAACAAG  
 AACCAGTCTG CGCTGTCGAG TTCTATCGAG CGTCTGTCTT CTGGCTTGCG TATTAACAGC  
 GCGAAGGATG ACGCCGCGAG TCAGGCGATT GCTAACCGTT TTACTTTCAA CATTAAAGGC  
 CTGACTCAGG CGGCCCGTAA CGCCAACGAC GGTATTCTGT TTGCGCAGAC CACCGAAGGC  
 GCGCTGTCCG AAATCAACAA CAACTTACAG CGTATTGCTG AACTGACGGT TCAGGCCACT  
 ACAGGGACTA ACTCCGATTG TGACCTGGAC TCCATCCAGG ACGAAATCAA ATCTCGTCTT  
 GATGAAATTG ACCGCGTATC CGGCCAGACC CAGTTCAACG GCGTGAACGT GCTGGCGAAA  
 GACGGTTCAA TGAAAAATCA GGTGGGTGCG AATGACGGCG AAACCATCAC GATCGACCTG  
 AAAAAATCG ATTCTGATAC TCTGGGTCTG AATGGCTTTA ACGTAAATGG TAAAGGTACT  
 ATTACCAACA AAGCTGCAAC GGTAAAGTAT TTAACCTCTG CTGGCGCGAA GTTAAACAC  
 CACGACAGGT CTTTATGATC TGAAAACCGA AAATACCTTG TTAACCTACG ATGCTGCATT  
 CGATAAATTA GGGAAATGGCG ATAAAGTCAC AGTTGGCGGC GTAGATTATA CTTACAACGC  
 TAAATCTGGT GATTTTACTA CCACTAAATC TACTGCTGGT ACGGGTGTAG ACGCCGCGGC  
 GCAGGCTGCT GATTGAGCTT CAAAACGTGA TCGTTAGCT GCCACCCCTC ATGCTGATGT  
 GGGTAAATCT GTTAAATGGT CTTACACCAC AAAAGATGGT ACTGTTTCTT TCGAAAACGGA  
 TTCAGCAGGT AATATCACCA TCGTGGAAG CCAGGCATAC GTAGACGATG CAGGCAACTT  
 GACGACTAAC AACGCTGGTA GCGCAGCTAA AGCTGATATG AAAGCGCTGC TCAAGCAGC  
 GAGCGAAGGT AGTGACGGTG CCTCTGTGAC ATTCAATGGC ACAGAATATA CCATCGCAAA  
 AGCAACTCCT GCGACAACCA CTCAGTAGC TCCGTTAATC CCTGGTGGGA TTACTTATCA  
 GGTACAGTG AGTAAAGATG TAGTATTGAG CGAAACCAAA CGCGCTGCGG CGACATCTTC  
 AATTACCTTT AATCCGGTG TACTGAGCAA AACTATTGGG TTTACCGCGG GTGAATCCAG  
 TGATGCTGCG AAGTCTTATG TGGATGATAA AGGTGGTATC ACTAACCGTT CCGACTATAC  
 AGTCTCTTAC AGCGTTAACA AGGATAACGG CTCTGTGACT GTTGCCGGGT ATGCTTCAGC  
 GACTGATACC AATAAAGATT ATGCTCCAGC AATTGGTACT GCTGTAATGT TGAACCTCCG  
 GGGTAAATC ACTACTGAGA CTACCACTGC TGGTTCTGCA ACGACCAACC CGCTTGCTGC  
 CTGTGGACGAC GCAATCAGCT CCATCGACAA ATTCCGTTCT TCCTGGGTG CTATCCAGAA  
 CCGTCTGGAT TCCGCACTCA CCAACCTGAA CAACACCACT ACCAACCTGT CCGAAGCGCA  
 GTCCTGTATT CAGGACGCGG ACTATGCGAC CGAAGTGTC AACATGTCGA AAGCGCAGAT  
 CATTCCAGAG GCCGGTAACT CCGTGCTGGC AAAAGCTAAC CAGGTACCGC AGCAGGTTCT  
 GTCTCTGCTG CAGGGTTAA

Figure 60



83/96

ATGGCACAAAG TCATTAATAC CAACAGCCTC TCGCTGATCA CTCAAAATAA TATCAACAAG  
 AACCAAGTCTG CGCTGTCGAG TTCTATCGAG CGTCTGTCTT CTGGCTTGCG TATTAACAGC  
 CGGAAGGATG ACGCAGCGGG TCAAGCGATT GCTAACCGTT TTACTTCTAA CATTAAAGGC  
 CTGACTCAGG CTGCACGTAA CGCCAACGAC GGTATTTCGT TTGCGCAGAC CACCGAAGGC  
 GCGCTGTCCG AATCAACAA CAACCTACAG CGTATTCTGT AACTGACGGT TCAGGCCACT  
 ACAGGGACTA ACTCCGATTC TGACCTGGAG TCCATCCAGG ACGAAATCAA ATCTCGTCTT  
 GATGAAATTTG ACCGCGTATC CGGCCAGACC CAGTTCAACG GCGTGAACGT GCTGGCGAAA  
 GACGGTTCAA TGAAATTTCA GGTGGGTGCG AATGACGGCG AAACCATCAC GATCGACCTG  
 AAAAAATCG ATTCTGATAC TCTGGGTCTG AATGGCTTTA ACGTAATGG TAAAGTACT  
 ATTACCAACA AAGCTGCAAC GGTAAGTGAT TTAACCTCTG CTGGCGCGAA GTTAAACAC  
 CACGACAGGT CTTTATGATC TGAAAACCGA AAATACCTTG TTAACACCG ATGCTGCAAT  
 CGATAAATTA GGGAAATGGCG ATAAAGTCAC AGTTGGCGGC GTAGATTATA CTTACAACGC  
 TAAATCTGGT GATTTTACTA CCACTAAATC TACTGCTGGT ACGGGTGTAA ACGCGCGCGC  
 GCAGGCTGCT GATTACGCTT CAAAACGTGA TGCCTTAGCT GCCACCTTC ATGCTGATGT  
 GGGTAAATCT GTTAATGGTT CTTACACCAC AAAAGATGGT ACTGTTTCTT TCGAAAACGGA  
 TTCAGCAGGT AATATCACC TCGGTGGAAG CCAGGCATAC GTAGACGAT CAGGCAACTT  
 GACGACTAAC AACGCTGGTA GCGCAGCTAA AGCTGATATG AAAGCGCTGC TCAAAGCAGC  
 GAGCGAAGGT AGTGACGGTG CCTCTCTGAC ATTCAATGGC ACAGAATATA CCATCGCAAA  
 AGCAACTCTT GCGACAACCA CTCAGTAGC TCCGTTAATC CCGGTGGGA TTACTTATCA  
 GGCTACAGTG AGTAAAGATG TAGTATTGAG CGAAACCAAA GCGGCTGCCG CGACATCTTC  
 AATTACCTTT AATCCGGTG TACTGAGCAA AACTATTGGG TTACCGCGG GTGAATCCAG  
 TGATGCTGCG AAGTCTTATG TGGATGATAA AGGTGGTATC ACTAACGTTG CCGACTATAC  
 AGTCTCTTAC AGCGTTAACA AGGATAACGG CTCTGTGACT GTTGCCGGGT ATGCTTCAGC  
 GACTGATACC AATAAAGATT ATGCTCCAGC AATTGGCACT GCTGTAAATG TGAACCTCCG  
 GGGTAAATC ACTACTGAGA CTACCAAGTC TGGTTCTGCA ACGACCAACC CGCTTGCTGC  
 CCTGGACGAC GCAATCAGCT CCATCGACAA ATTCCGTTCT TCCTGGGTG CTATCCAGAA  
 CCGTCTGGAT TCCGCGGTCA CCAACCTGAA CAACACCACT ACCAACCTGT CCGAAGCGCA  
 GTCCCGTATT CAGGACGCGC ACTATGCGAC CGAAGTGTC AACATGTGCA AAGCGCAGAT  
 CATCCAGCAG GCGGTAACT CCGTGCTGGC AAAAGCTAAC CAGGTACCGC AGCAGGTTCT  
 GTCTCTGCTG CAGGTTTAA

Figure 61

84/96

ATGGCACAAG TCATTAATAC CAACAGCCTC TCGTGATCA CTCAAAAAT TATCAACAAG  
 AACCAAGCTG CGCTGTCGAG TTCTATCGAG CGTCTGTCTT CTGGCTTGCG TATTAACAGC  
 GCGAAGGATG ACGCCGCGGG TCAGGCGATT GCTAACCGTT TTACTTCTAA CATTAAAGGC  
 CTGACTCAGG CTGCACGTAA CGCCAACGAC GGTATTTCTG TTGCACAGAC CACTGAAGGC  
 GCGCTGTCCG AAATCAACAA CAACTTACAG CGTATCCGTG AGCTGACGGT TCAGGCTTCT  
 ACCGGGACTA ACTCTGATTC GGATCTGGAC TCCATTCAAG ACGAAATCAA ATCCCGTCTC  
 ACGCAAAATT ACCGCGTATC CGGTCAAGAC CAGTCAACG CGGTGAACGT ACTGGCAAAA  
 GACGGTTCGA TGAATAATCA GGTGGTGGC AATGACGGTG AAATATCAC TATCGACCTG  
 AAGAAAAATC ATTCTGATAC TCTGGGTCTG AATGGTTTAA ACGTAAATGG TAAAGGTACT  
 ATTACCAACA AAGCTGCAAC GGTAAAGTAT TTAACCTCTG CTGGCGCGAA GTTAAACAC  
 CACGACAGGT CTTTATGATC TGAATAACGA AAATACCTTG TTAACACCG ATGCTGCATT  
 CGATAAATTA GGAATGGCG ATAAAGTCAC CGTTGGCGGC GTAGATTATA CTTACAACGC  
 TAAATCTGGT GATTTTACTA CCACCAATC TACTGCTGGT ACGGGTGTAG ACGCCGCGGC  
 GCAGGCTACT GATTACGCTA AAAAACGTGA TGGCTTAGCT GCCACCCCTC ATGCTGATGT  
 GGGTAAATCT GTTAATGGTT CTTACACCAC AAAAGATGGT ACTGTTTCTT TCGAAACCGA  
 TTCAGCAGT AATATACCA TCGGTGGAAG CCAGGCATAC GTAGACGAT CAGGCACTT  
 GACGACTAAC AACGCTGGTA GCGCAGCTAA AGCTGATATG AAAGCGCTGC TTAAGCCGCG  
 GAGCGAAGGT AGTGACGGTG CCTCTCTGAC ATTCAATGGC ACTGAATATA CTATCGCAAA  
 AGCAACTCCT GCGACAACCT CTCCAGTAGC TCCGTTAATC CCTGGTGGGA TTACTTTATCA  
 GGCTACAGTG AGTAAAGATG TAGTATTGAG CGAAACCAAA GCGGCTGCCG CGACATCTTC  
 AATTACCTTT AATCCGGTG TACTGAGCAA AACTATTGGG TTTACCGCGG GTGAATCCAG  
 TGATGCTGCG AAGTCTTATG TGGATGATAA AGGTGGTATT ACTAACGTTG CGACTATAC  
 AGTCTCTTAC AGCGTTAACA AGGATAACGG CTCTGTGACT GTTGCCGGGT ATGCTTCAGC  
 GACTGATACC AATAAAGATT ATGCTCCAGC AATTGGTACT GCTGTAAATG TGAACCTCCG  
 GGGTAAATC ACTACTGAGA CTACCAAGTGC TGGTTCTGCA ACGACCAACC CGCTTGTCTG  
 CCTGGACGAC GGTATCAGT CCATCGACAA ATTCCGTCTT TCCCTGGGTG CTATCCAGAA  
 CCGTCTGGAT TCCGAGTCA CCAACCTGAA CAACACCACT ACCAACCTGT CTGAAGCGCA  
 GTCCCGTATT CAGGACGCCG ACTATGCGAC CGAAGTGTCC AACATGTCGA AAGCGCAGAT  
 TATCCAGCAG GCCGGTAACT CCGTGTCTGC AAAAGCCAAC CAGGTACCGC AGCAGGTTCT  
 GTCTCTGCTG CAGGGTTAA

Figure 62

85/96

ATGGCACAAGTCATTAATACCAACAGCCTCTCGCTGATCACTCAAATAATATCAACAAG  
 AACCAAGTCTGCGCTGTCGAGTTCTATCGAGCGTCTGTCTTCTGGCTTSCGTATTAAACAGC  
 GCGAAGGATGACGCGCAGGTACGCGGATTGTAAACGTTTTACTTCTAACTTAAAGGC  
 CTGACTCAGGCGGCCCCGTAACGCCAACGCGGTATTTCTGTTGCGCAGACCACCGAAGGC  
 GCGCTGTCCGAAATCAACAACACTTACAGCGTATTCTGTAACTGACGGTTACAGGCCACT  
 ACAGGGACTAACTCCGATTCTGACCTGGACTCCATCCAGGACGAAATCAAATCTGCTCTT  
 GATGAAATTGACCGCGTATCCGGCCAGACCCAGTTCAACGCGGTGAACGTGCTGGCCGAAA  
 GACGTTCAATGAAATTCAGGTTGGTGCGAATGACGCGAAACCATCACGATCGACCTG  
 AAAAAATCGATTCTGATACTCTGGGTCTGAATGGCTTAACTGAATGGTAAAGGTACT  
 ATTACCAACAAAGCTGCAACGGTAAAGTATTAACTTCTGCTGGCGCGAAGTTAAACACC  
 ACGACAGGCTCTTATGATCTGAAAACCGAAAACTCTTGTAACTACCGATGCTGCATT  
 GATAAATTAGGGAATGGCGATAAAGTCAAGTTGGCGCGTAGATTATACCTACAACGCT  
 AAATCTGGTGATTTTACTACCACTAAATCTACTGCTGGTACGGGTGACGCGCGCGGGC  
 CAGGCTGCTGATTACGCTTCAAAACGTGATGCGTTAGCTGCCACCCCTTCATGCTGATGTG  
 GGTAATCTGTTAATGGTTCTTACACCACAAAGATGGTACTGTTCTTTCGAAAACGGAT  
 TCAGCAGGTAATATCACCATCGGTGGAAGCCAGGCATACGTAGACGATGCAGGCAACTTG  
 ACGACTAACACGCTGGTAGCGCAGCTAAAGCTGATATGAAAGCGCTGCTCAAAGCAGCG  
 AGCGAAGGTAGTGACGGTGCCTCTCTGACATTCAATGGCACAGAATATACCATCGCAAAA  
 GCAACTCCTGCGACAACCACTCCAGTAGCTCCGTTAATCCCTGGTGGATTACTTATCAG  
 GCTACAGTGAGTAAAGATGTAGTATTGAGCGAAACCAAAGCGGGTGCCTGCGCATCTTCA  
 ATTACCTTTAATCCGGTGTACTGAGCAAACTATTGGGTTTACCGCGGGTGAATCCAGT  
 GATGCTGCGAAGTCTTATGTGGATGATAAAGGTGGTATCACTAACGTTTGCCTACTATACA  
 GTCTCTTACAGCGTTAACAAGGATAACGGCTCTGTGACTGTTGCGCGGTATGCTTCAGCG  
 ACTGATACCAATAAAGATTATGCTCCAGCAATTTGGTACTGCTGTAAATGTGAACCTCCGCG  
 GGTAAAACTCACTACTGAGACTACCAAGTCTGGTTCTGCAACGACCAACCCGCTTGTGTC  
 CTGGACGACGCAATCAGTCCATCGACAAATCCGTTCTTCCCTGGGTGCTATCCAGAAC  
 CGTCTGGATTCCGCGAGTCAACCACTGAACCAACCACTACCACTGTCCGAAGCGCAG  
 TCCCGTATTAGGACGCGAGCTATGCGACCGAAGTGTCCAACTGTGAAAAGCGCAGATC  
 ATTACGAGGCGGTAACTCCGTGCTGGCAAAAGCTAACCAAGTACCGCAGCAGGTTCTG  
 TCTCTGCTGCAGGTTAA

Figure 63

ATGGCACAAG TCATTAATAC CAACAGCCTC TCGCTGATCA CTCAAAATAA TATCAACAAG  
 AACCAAGTCTG CGCTGTCGAG TTCATATCGAG CGTCTGCTCT CTGGCTTGGC TATTAACAGC  
 GCGAAGGATG ACGCCGCGAGG TCAGCGGATT GCTAACCGTT TTACTTCTAA CATTAAAGGC  
 CTGACTCAGG CTGCACGTAA CGCCAACGAC GGTATTTCTG TTGCGCAGAC CACCGAAGGC  
 GCGCTGTCG AAATCAACAA CAACTTACAG CGTATTCGTG AACTGACGGT TCAGGCCACT  
 GATGAAATTG ACCGCGTATC CGGCCAGACC CAGTTCAACG GCGTGAACGT GCTGGCGAAA  
 GACGGTTCAA TGAAAAATCA GGTGGTGCG AATGACGGCG AAACCATCAC GATCGACCTG  
 AAAAAATCG ATTCTGATAC TCTGGGTCGT AATGGCTTTA ACGTAAATGG TAAAGGTACT  
 ATTACCAACA AAGCTGCAAC GGTAAAGTAT TTAACCTCTG CTGGCGCGAA GTTAAACAC  
 CACGACAGGT CTTTATGATC TGAAAACCGA AAATACCTTG TTAACACCG ATGCTGCATT  
 CGATAAATTA GGAATGGCG ATAAAGTCAC AGTTGGCGGC GTAGATTATA CTTACACGC  
 TAAATCTGGT GATTTTACTA CCACTAAATC TACTGCTGGT ACGGGTGTAG ACGCCGCGGC  
 GCAGGCTGCT GATTCAAGCTT CAAAACGTGA TCGGTTAGCT GCCACCTTTC ATGCTGATGT  
 GGGTAAATCT GTTAATGGTT CTTACACCAC AAAAGATTGGT ACTGTTTCTT TCGAAACGGA  
 TTCAGCAGT AATATCACC TCGGTGGAAG CCAGGCATAC GTAGACGATG CAGGCAACTT  
 GACGACTAAC AAGCTGGTA GCGCAGCTAA AGCTGATATG AAAGCGCTGC TCAAAGCAGC  
 GAGCGAAGGT AGTGACGGT CCTCTCTGAC ATTCAATGGC ACAGAAATATA CCATCGCAAA  
 AGCAACTCCT GCGACAACCA CTCCAGTAGC TCCGTTAATC CCTGGTGGGA TTACTTATCA  
 GGCTACAGTG AGTAAAGATG TAGTATTGAG CGAAACCAAA GCGGCTGCCG CGACATCTTC  
 AATTACCTTT AATTCCGGTG TACTGAGCAA AACTATTGGG TTTACCGCG GTGAATCCAG  
 TGATGCTGCG AAGTCTTATG TGGATGATAA AGGTGGTATC ACTAACGTTG CCGACTATAC  
 AGTCTCTTAC AGCGTTAACA AGGATAACGG CTCTGTGACT GTTGCCGGGT ATGCTTCAGC  
 GACTGATACC AATAAGATT ATGCTCCAGC AATTGGCACT GCTGTAAATG TGAACCTCCG  
 GGGTAAATC ACTACTGAGA CTACCAAGTG TGGTTCTGCA ACGACCAACC CGCTTGCTGC  
 CCTGGACAGC GCAATCAGCT CCATCGACAA ATTCCGTTCT TCCCTGGGTG CTATCCAGAA  
 CGTCTGGAT TCCGCGGTCA CCAACCTGAA CAACACCACT ACCAACCTGT CCGAAGCGCA  
 GTCCCGTATT CAGGACGCCG ACTATGCGAC CGAAGTGTC AACATGTCGA AAGCGCAGAT  
 CATCCAGCAG GCCGGTAACT CCGTGCTGGC AAAAGCTAAC CAGGTACCGC AGCAGGTTCT  
 GTCTCTGCTG CAGGGTTAA

Figure 64

87/96

ATGGCACAAG TCATTAATAC CAACAGCCTC TCGCTGATCA CTCAAAATAA TATCAACAG  
 AACCAAGTCTG CGCTGTGCGAG TTCTATCGAG CGTCTGTCTT CTGGCTTGGC TATTAACAGC  
 GCGAAGGATG ACGCCGCGGG TCAGGCGATT GCTAACCGTT TTACTTCTAA CATTAAAGGC  
 CTGACTCAGG CTGCACGTAA CGCCAACGAC GGTATTCTG TTGCACAGAC CACTGAAGGC  
 CGCGTGTCCG AAATCAACAA CACTTACAG CGTATCCGTG AGCTGACGGT TCAGGCTCTC  
 ACCGGGACTA ACTCTGATTC GGATCTGGAC TCCATTGAGG ACGAAATCAA ATCCCGTCTC  
 GACGAAATTG ACCGCGTATC CGGTGAGACC CAGTTCACG GCGTGAACGT ACTGGCAAAA  
 GACGGTTCGA TGAAAAATTCA GGTGGTGGC AATGACGGTG AAATATCAC TATCGACCTG  
 AAGAAAATCG ATTCTGATAC TCTGGGCTG AATGGTTTTA ACGTAAATGG TAAAGGTACT  
 ATTACCAACA AAGCTGCAAC GGTAAGTGAT TTAACCTCTG CTGGCGCGAA GTTAAACACC  
 ACGACAGGT CTTTATGATC TGAAAAACGA AAATACCTTG TTAACACCG ATGCTGCATT  
 CGATAAATTA GGAATGGCG ATAAAGTCAC CGTTGGCGGC GTAGATTATA CTTACAACGC  
 TAAATCTGGT GATTTTACTA CCACCAATC TACTGCTGGT ACGGGTGTAG ACGCCGCGGC  
 GCAGGCTACT GATTGAGCTA AAAAACGTGA TCGGTTAGCT GCCACCTTC ATGCTGATGT  
 GGGTAAATCT GTTAATGGTT CTTACCCAC AAAAGATGGT ACTGTTTCTT TCGAAACGGA  
 TTCAGCAGGT AATATCACCA TCGGTGGAAG CCAGGCATAC GTAGACGATG CAGGCAACTT  
 GACGACTAAC AACGCTGGTA GCGCAGCTAA AGCTGATATG AAAGCGCTGC TTAAGCCGCG  
 GAGCGAAGGT AGTGACGGTG CCTCTCTGAC ATTCAATGGC ACTGAATATA CTATCGCAAA  
 AGCAACTCTCT GCGACAACCT CTCCAGTAGC TCCGTTAATC CCTGGTGGGA TTTCTTATCA  
 GGCTACAGTG AGTAAAGATG TAGTATTGAG CGAAACCAAA GCGGCTGCCG CGACATCTTC  
 AATTACCTTT AATTCGGTG TACTGAGCAA AACTATTGGG TTTACCGCGG GTGAATCCAG  
 TGATGCTGCG AAGTCTTATG TGGATGATAA AGGTGGTATT ACTAACGTTG CCGACTATAC  
 AGTCTCTTAC AGCGTTAACA AGGATAACGG CTCTGTGACT GTTGCCGGGT ATGCTTCAGC  
 GACTGATACC AATAAGATT ATGCTCCAGC AATTGGTACT GCTGTAATG TGAACCTCCG  
 GGGTAAATC ACTACTGAGA CTACCACTGC TGGTTCTGCA ACGACCAACC CGCTTGCTGC  
 CCTGGACGAC GCTATCAGCT CCAATCGCAA ATTCGGTCTC TCCTGGGTG CTATCCAGAA  
 CGGCTGGAT TCCGCACTCA CCAACCTGAA CAACACCACT ACCAACCTGT CTGAAGCGCA  
 GTCCCGTATT CAGGACGCCG ACTATGCGAC CGAAGTGCC AACATGTGCA AAGCGCAGAT  
 TATCCAGCAG GCCGTAACT CCGTGCTGGC AAAAGCCAA CAGGTACCGC AGCAGGTTCT  
 GTCTCTGCTG CAGGGTTAA

Figure 65

88/96

ATGGCACAAG TCATTAATAC CAACAGCCTC TCGCTGATCA CTCAAAATAA TATCAACAAG  
 AACCAAGTCTG CGCTGTCGAG TTCTATCGAG CGTCTGTCTT CTGGCTTGGG TATTAAACAGC  
 GCGAAGGATG ACGCCGCGAGG TCAGGCGATT GCTAACCGTT TTACTTCTAA CATTAAAGGC  
 CTGACTCAGG CGGCCCGTAA CGCCAACGAC GGTATTCTGT TTGCGCAGAG CACCGAAGGC  
 GCGCTGTCCG AAATCAACAA CAACTTACAG CGTATTCTGT AACTGACGGT TCAGGCCACT  
 ACAGGGACTA ACTCCGATTG TGACCTGGAC TCCATCCAGG ACGAAATCAA ATCTCGTCTT  
 GGTGAAATTG ACCGCGTATC CGGCCAGACC CAGTTCAACG GCGTGAACGT GCTGGCGAAA  
 GACGGTTCAA TGAAAATTCA GGTGTGTGCG AATGACGGCG AAACCATCAC GATCGACCTG  
 AAAAAAATCG ATTCTGATAC TCTGGGTCTG AATGGCTTTA ACGTAAATGG TAAAGGTACT  
 ATTACCAACA AAGCTGCAAC GGTAAGTGAT TTAACCTCTG CTGGCGCGAA GTTAAACAC  
 CACGACAGGT CTTTATGATC TGAAAACCGA AAATACCTTG TTAACCTACG ATGCTGCATT  
 CGATAAATTA GGAATGGCG ATAAAGTCAC AGTTGGCGGC GTAGATTATA CTTACAACGC  
 TAAATCTGGT GATTTTACTA CCACTAAATC TACTGCTGGT ACGGCTGTAG ACGCCGCGGC  
 GCAGGCTGCT GATTCAGCTT CAAAACGTGA TGCCTTAGCT GCCACCCTTC ATGCTGATGT  
 GGGTAAATCT GTTAATGGTT CTTACACCAC AAAAGATGGT ACTGTTTCTT TCGAAACGGA  
 TTCAGCAGGT AATATCACCA TCGGTGGAAG CCAGGCATAC GTAGACGATG CAGGCAACTT  
 GACGACTAAC AACGCTGGTA GCGCAGCTAA AGCTGATATG AAAGCGCTGC TCAAAGCAGC  
 GAGCGAAGGT AGTGACGGTG COTCTCTGAC ATTCATGGC ACAGAATATA CCATCGCAAA  
 AGCAACTCCT GCGACAACCA CTCAGTAGC TCCGTAAATC CCTGGTGGGA TTACTTATCA  
 GGCTACAGTG AGTAAAGATG TAGTATTGAG CGAAACCAAA GCGGGTGGCG CGACATCTTC  
 AATTACCTTT AATTCGGGTG TACTGAGCAA AACTATTGGG TTACCGCGG GTGAATCCAG  
 TGATGCTGCG AAGTCTTATG TGGATGATAA AGGTGGTATC ACTAACGTTG CCGACTATAC  
 AGTCTCTTAC AGCGTTAACA AGGATAACGG CTCTGTGACT GTTGCCGGGT ATGCTTCAGC  
 GACTGATACC AATAAGATT ATGCTCCAGC AATTGGTACT GCTGTAAATG TGAATCCGC  
 GGGTAAATC ACTACTGAGA CTACCAGTGC TGGTCTGCGA ACGACCAACC CGCTTGTCTG  
 CCTGGACGAC GCAATCAGCT CCATCGACAA ATTCCGTCTT TCCCTGGGTG CTATCCAGAA  
 CCGTCTGGAT TCCGCACTCA CCAACCTGAA CAACACCACT ACCAACCTGT CCGAAGCGCA  
 GTCCCGTATT CAGGACGCCG ACTATGCGAC CGAAGTGTC AACTGTGCGA AAGCGCAGAT  
 CATTACAGC GCGGTAACCT CCGTGTGCG AAAAGCTAAC CAGGTACCGC AGCAGGTCTT  
 GTCTCTGCTG CAGGGTTAA

Figure 66

89/96

ATGGCACAAAG TCATTAATAC CAACAGCCTC TCGCTGATCA CTCAAAATAA TATCAACAAG  
 AACCAAGCTG CGCTGTCGAG TTCTATCGAG CGTCTGCTT CTGGCTTGGC TATTAACAGC  
 GCGAAGGATG ACGCCGCGGG TCAGGCGGATT GCTAACCGTT TTACTTCTAA CATTAAAGGC  
 CTGACTCAGG CTGCACGTAA CGCCAACGAC GSTATTCTG TTGCACAGAC CACCGAAGGC  
 GCGCTGCTG AAATCAACAA CAACTTACAG GGTATCCGTG AGCTGACGGT TCAGGCTTCT  
 ACCGGAACCTA ACTCTGATTG GGATCTGGAC TCCATTTCAGG ACGAAATCAA ATCCCGTCTT  
 GATGAAATG ACCGCGTATC CGGCCAGACC CAGTTCAACG GCGTGAACGT ACTGGCAAAA  
 GACGGTTCGA TGAAAATTCA GGTGGTGGC AATGACGGT AACTATCAC TATCGACCTG  
 AAGAAAAATG ATTCTGATAC TCTGGTCTG AATGGTTTA ACGTAAATGG TAAAGGTACT  
 ATTACCAACA AAGCTGCAAC GGTAAAGTAT TTAATTCTG CTGGCGCGAA GTTAAACAC  
 CACGACAGGT CTTTATGATC TGAAAACCGA AAATACCTTG TTAACACCG ATGCTGCATT  
 CGATAAATTA GGGAAATGGC ATAAAGTCAC CGTTGGCGGC GTAGATTATA CTTACAAACG  
 TAAATCTGGT GATTTTACTA CCACCAATC TACTGCTGGT ACGGGTGTAG ACGCCGCGGC  
 GCAGGCTACT GATTGAGCTA AAAAACGTGA TCGTTAGCT GCCACCTTC ATGCTGATGT  
 GGGTAAATCT GTTAATGGTT CTTACACCAC AAAAGATGGT ACTGTTTCT TCGAAACGGA  
 TTCAGCAGGT AATATCACCA TCGGTGGAAG CACGGCATACT GTAGACGATG CAGGCAACTT  
 GACGACTAAC AACGCTGGTA GCGCAGCTAA AGCTGATATG AAAGCGCTGC TTAAGCCGCG  
 GAGCGAAGGT AGTGACGGTG CTCTCTGAC ATTCAATGGC ACTGAATATA CTATCGCAAA  
 AGCAACTCCT GCGACAACT CTCCAGTAGC TCCGTTAATC CTTGGTGGGA TTACTTATCA  
 GGCTACAGTG AGTAAAGATG TAGTATTGAG CGAAACCAAA GCGGCTGCGG CGACATCTTC  
 AATTACCTTT AATCCGGTG TACTGAGCAA AACTATTGGG TTTACCGCGG GTGAATCCAG  
 TGATGCTCGG AAGTCTTATG TGGATGATAA AGGTGGTATT ACTAACGTTG CCGACTATAC  
 AGTCTCTTAC AGCGTTAACA AGGATAACGG CTCTGTGACT GTTGCCGGGT ATGCTTCAGC  
 GACTGATACC AATAAAGATT ATGCTCCAGC AATTGGTACT GCTGTAATG TGAACCTCCG  
 GGGTAAATC ACTACTGAGA CTACCAAGTGC TGGTTCTGCA ACGACCAACC CGCTTGTGCG  
 CCTGGACGAC GCTATCAGCT CCATCGACAA ATCCGTTCT TCCCTGGGTG CTATCCAGAA  
 CCGTCTGGAT TCCGAGTCA CCAACCTGAA CAACACCACT ACCAACCTGT CTGAAGCGCA  
 GTCCCGTATT CAGGACGCG ACTATGCGAC CGAAGTGTCC AACATGTCGA AAGCGCAGAT  
 TATCCAGCAG GCCGGTAACT CCGTGTGGC AAAAGCCAAC CAGGTACCGC AGCAGGTTCT  
 GTCTCTGCTG CAGGGTTAA

Figure 67

90/96

ATGGCACAAAG TCATTAATAC CAACAGCCTC TCGTGATCA CTCAAATAA TATCAACAAG  
 AACCAAGCTCG CGCTGTCGAG TTCTATCGAG CGTCTGTCTT CTGGCTTGGG TATTAACAGC  
 GCGAAGGATG ACGCCGCGGG TCAGGCGATT GCTAACCGTT TTACTTCTAA CATTAAAGGC  
 CTGACTCAGG CTGCACGTAA CGCCAACGAC GGTATTCTGT TTGCACAGAC CACTGAAGGC  
 GCGCTGTCCG AAATCAACAA CAACTTACAG CGTATCCGTG AGCTGACGGT TCAGGCTTCT  
 ACCGGGACTA ACTCTGATTG GGATCTGGAC TCCATTGAG ACGAATCAA ATCCCGTCTC  
 GACGAATATG ACCGCGTATC CGGTCAGACC CAGTTCAACG GCGTGAACGT ACTGGCAAAA  
 GACGGTTCTGA TGAATAATCA GGTGTGGTGG AATGACGGTG AAATATCAC TATCGACCTG  
 AAGAAAATCG ATTCTGATAC TCTGGGTCTG AATGGTTTTA ACGTAAATGG TAAAGGTACT  
 ATTACCAACA AAGCTGCAAC GGTAAGTGAT TTAACTTCTG CTGGCGCGAA GTTAAACAC  
 CACGACAGGT CTTTATGATC TGAAAACCGA AAATACCTTG TTAACCTACG ATGCTGCATT  
 CGATAAATTA GGAATGGCG ATAAAGTCAC CGTTGGCGGC GTAGATTATA CTTACAACGC  
 TAAATCTGGT GATTTTACTA CCACCAAATC TACTGCTGGT ACGGGTGTAG ACGCCGCGGC  
 GCAGGCTACT GATTGAGCTA AAAAACGTGA TGCCTTAGCT GCCACCCCTC ATGCTGATGT  
 GGTAATATCT GTTAATGGTT CTTACACCAC AAAAGATGGT ACTGTTTCTT TCGAAACGGA  
 TTCAGCAGGT AATATCACCA TCGGTGGAAG CCAGGCATAC GTAGACGATG CAGGCACTT  
 GACGACTAAC AACGCTGGTA GCGCAGCTAA AGCTGATATG AAAGCGCTGC TTAAGCCGC  
 GAGCGAAGGT AGTGACGGTG CCTCTGTGAC ATTCAATGGC ACTGAATATA CTATCGCAAA  
 AGCAACTCCT GCGACAACCT CTCCAGTAGC TCCGTTAATC CTGGGTGGGA TTTCCTATCA  
 GGCTACAGTG AGTAAAGATG TAGTATTGAG CGAAACCAA GCGGCTGCCG CGACATCTTC  
 AATTACCTTT AATTCCGGTG TACTGAGCAA AACTATTGGG TTTACCGCGG GTGAATCCAG  
 TGATGCTGGC AAGTCTTATG TGGATGATAA AGSTGGTATT ACTAACGTTG CCGACTATAC  
 AGTCTCTTAC AGCGTTAACA AGGATAACGG CTCTGTGACT GTTGCCGGGT ATGCTTCAGC  
 GACTGATACC AATAAGATT ATGCTCCAGC AATTGGTACT GCTGTAAATG TGAACCTCGC  
 GGGTAAATC ACTACTGAGA CTACCAGTGC TGGTTCTGCA ACGACCAACC CGCTTGCTGC  
 CCTGGACGAC GCTATCAGCT CCATCGACAA ATTCCGTTCT TCCCTGGGTG CTATCCAGAA  
 CCGTCTGGAT TCCGAGTCA CCAACCTGAA CAACCCACT ACCAACCTGT CTGAAGCGCA  
 GTCCCGTATT CAGGACCGCG ACTATGCGAC CGAAGTGTCC AACATGTGCA AAGCGCAGAT  
 TATCCAGCAG GCCGGTAACT CCGTGTGGC AAAAGCCAAC CAGGTACCGC AGCAGGTTCT  
 GTCTCTGCTG CAGGGTTAA

Figure 68



ATGCGACGTATAGAACGAATACCGGGGTTATCGGCGTAAGCGGGGCAAA  
 GTTTACGATTTATTTTTGGGCTTAATGACACGAACAGCAACGAGGAAGGG  
 GAGTATTTTCGACCGCTAGAAAAAATTCTAAAGGTTGTGAGTGACCAGAC  
 GATAACAGGGTTGACGGCGACGAAGCCGAAGGGTGAAGCCCAATACTT  
 AAACCGTAGACTTGAAAAACAGGAAAAATGAATCATGGCACAAGTCATTAAT  
 ACCAACAGCCTCTCGCTGATCACTCAAAATAATATCAACAAGAACCGTCT  
 TGGCTGTGCGACTTCTATCGAGCGCTCTCTTCTGGTCTGCGCATTAACAG  
 CGCTAAAGATGACGCTGCGGGCCAAGCGATTGCTAACCGCTTCACTTCTA  
 ACATCAAAGGTCTGACTCAGGCGCACGTAACGCCAACGACGGATTTCT  
 CTGGCGCAGACCACTGAAGGCGCACTGTCTGAAATCAACAACAACCTTGCA  
 GCGTGTTCGTGAACCTGACCGTTCAGGCCACTACCGGTAACCTCTGATTCT  
 TGACCTGTCTTCAATACAGGACGAAATCAAATCCCGTCTCGATGAAATTG  
 ACCGCGTATCCGGTCAGACTCAGTTCAACGGCGTTAATGTTCTTTCCAAAG  
 ATGGTTCAATGAAAAATTCAGGTTGGTGCGAATGATGGTCAAACTATCTCC  
 ATCGATCTGAAGAAAATTGATTCTTCAACTTTGGGGCTGAATGGCTTCTCA  
 GTTTCTAAAAAATCTCTTAATGTGACGAATGCTATCACATCTATCCCGCAA  
 GCCGCTAGCAATGAACCTGTTGATGTTAACTTCGGTGATACTGATGAGTCT  
 GCAGCAATCGCAGCCAAAATTGGGGTTTCCGATACGTCAAGCCTGTGCGCT  
 GCACAACATCCTTGATAAAGATGGTAAGGCAACAGCTGATTATGTTGTC  
 AGTCAGGTAAGACTTCTATGCTGCTTCTGTTAATGCCGCTTCAGGTAAAG  
 TAACCTTAAACACCAATTGATGTTACTTATGATGATTATGCGAACGGTGTG  
 ACGATGCCAAGCAACAGGTCAGCTGATCAAAGTTTCAGCAGATAAAGAC  
 GCGCAGCTCAAGGTTTTGTCACTTCAAGGCAAAAATATTCTGCTGGT  
 GATGCGGCAGACATTCTTAAGAATGGAGCAACAGCTCTTAAGTTAACTGA  
 TCTGAATTTAAGTGATGTTACTGATACTAATGGTAAGGTAACCAACAACCTGC  
 GACTGAGCAATTTGAAGGTGCTTCAACTGAGGATCCGCTGGCGCTTCTGG  
 ATAAAGCTATTGCATCAGTCGACAAATTCGGTCTTCTCTAGGTGCCGTGC  
 AGAACCGTCTCGATTCCGCTATCACAACCTGAACAACACCACCACCAAC  
 CTGTCTGAAGCGCAGTCCGTTATTCAGGACGCCGACTATGCGACCGAAGT  
 GTCCAACATGTGCAAGCGCAGATCATCCAGCAGGCAGGTAACTCCGTGC  
 TGTCTAAAGCGAACCAGGTACCGCAGCAAGTTCTGTCACTGTTACAAGGC  
 TAATGGCCTTAACTGCCTGACCCCGCCACCGGCGGGGTTTTTCTGTCCG  
 CAATTTACCGATAACCCCAATAACCCCTCATTTACCCCACTAATCGTCC  
 GATTAAAAACCCCTGCAGAAACGGATAATCATGCCGATAACCTATATAACG  
 CAGGGCTGTTTATCGTGAATTCCTCTATACCGCTGAAGGTGTAATGGATA  
 AACACTCGCTG

Figure 69

92/96

AACAGCCTCTGCTGATCACTCAGAACAAACATCAACAAAAACCAGTCTTC  
AATGTCTACTGCCATTGAGCGTCTGTCTTCCGGTCTGCGTATCAACAGCGC  
AAAAGATGACGCTGCTGGCCAGGCGATTGCCAACCGCTTCACCTCTAACA  
TCAAAGGTCTGACTCAGGCAGCTCGTAAACGCCAACGACGGTATCTCCGTT  
GCACAGACCACTGAAGGCGCACTGTCTGAAATCAACAACAACCTGCAGCG  
TATCCGTGAGCTGACTGTTCACTCTTCTACGGGTACTAACTCTGAATCCGA  
TCTGAACTCAATCCAGGACGAAATTAATCCCGTCTGGACGAAATTGACC  
GCGTATCCGGTCAGACCCAGTTCAACGGCGTGAAACGTGCTGGCAAAAGAC  
GGCTCCATGAAAAATTCAGGTTGGCGCGAACGATGGTGAAACCATCACCAT  
CGACCTGAAAAAAATTGACTCTTCTACTTTAAACCTGACTGGGTTTAA

Figure 70A

93/96

CTCAGTATGCTGTACCCGGCAGTACAGGTGCCGTAACCTACGATCCAGAT  
ACAGATCCTGCCGCGACTGGTGATATTGTTTCTGCTTATGTTGATGATGCA  
GGTACATTGACAACTGATGCAAAACAAAACGTAAAAATATTATGCCACAC  
TAATGGTAGCGTCACGAACGACAGTGGTTCAGCTATTTACGCAACTGAAG  
CGGGCAAATTGACTACTGAAGCGTCTACAGCTGCTGAAACTACCGCTAAC  
CCACTGAAAGCCCTGGACGATGCAATCAGCCAGATCGACAAATTCGGTTC  
TTCTCTGGGTGCTGTACAGAACCGTCTGGATTCTGCGGTAACCAACCTGAA  
CAACACCACCACCAACCTGTCTGAAGCGCAGTCCCGTATTCCAGGACGCCG  
ACTATGCGACCGAAGTGTCAAATATGTCTAAAGCGCAGATCATCCAGCAG  
GCCGGTAACTCCGTGTTGGCTAAAGCTAACCCAGGTTCTCAGCAGGTT

Figure 70B

94/96

AGCCTGTCGCTGTTGACCCAGAATAACCTGAACAAATCTCAGTCTTCTCTG  
AGCTCCGCCATTGAGCGTCTCTCTTCTGGCCTGCGTATTAACAGTGCTAAA  
GATGACGCAGCAGGTCAGGCGATTGCTAACCGTTTTACAGCAAATATTAA  
AGGTCTGACTCAGGCTTCCCGTAACGCGAATGATGGTATTTCTGTTGCGCA  
GACCACTGAAGGCGCGCTGAATGAAATTAACAACAACCTGCAGCGTGTA  
GTGAAC TGACTGTT CAGGCAACTAACGGTACTAACTCTGACAGCGATCTT  
CTTCTATCCAGGCTGAAATTACTCAACGCTCTGGAAGAAATTGACCGTGAT  
CTGAG CAAACTCAGTTTAACGGCGTGAAAGTCCTTGCTGAAAAAT

Figure 71

95/96

GCACGTTAGTTGTTAACGGTGCAACTTACGATGTTAGTGCAGATGGTAAA  
ACGATAACGGAGACTGCTTCTGGTAACAATAAAGTCATGTATCTGAGCAA  
ATCAGAAGGTGGTAGCCCCATTCTGGTAAACGAAGATGCAGCAAAATCGT  
TGCAATCTACCACCAACCGCTCGAAACTATCGACAAAAGCATTGGCTAAA  
GTTGACAACTCTGCGTTCTGACCTCGGTGCAGTACAAAACCGTTTCGACTCT  
GCTATCACCAACCTTGGCAACACCGTAAACAACCTGTCTCTGCCCGTAGC  
CGTATCGAAGATGCTGACTACGCGACCGAAGTGCTAACATGTCTCGTGC  
GCAGATCCTGCAACAAGCGGGTACCTCTGTTCTGGCGCAGGCTAACAGA  
CCACGCAGAACGTAC

Figure 72

96/96

FIGURE 73A

Sequence of the polylinker region of plasmid pTtc99A:

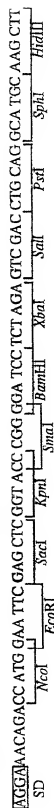
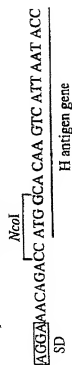


FIGURE 73B

Sequence in the junction region between vector and the 5' end of the H antigen gene:



## DECLARATION AND POWER OF ATTORNEY

As a below named inventor, I hereby declare:

That my residence, post office address and citizenship are as stated below next to my name.

That I verily believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural inventors are named below) of the subject matter which is claimed and for which a patent is sought on the invention entitled: **ANTIGENS AND THEIR DETECTION** the specification of which (check one)

☐ is attached hereto.

☒ was filed on 21/05/1999 as International Application Serial No. PCT/AU99/00385 and was amended on 01/05/2000 (if applicable).

That I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

That I acknowledge the duty to disclose information known to be material to patentability of this application in accordance with Title 37, Code of Federal Regulations §1.56(a).

That I hereby claim foreign priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate on this invention having a filing date before that of the application on which priority is claimed:

Prior Foreign Application(s)

PCT/AU99/00385	PCT	21/05/1999
(Number)	(Country)	(Day/Month/Year Filed)

Priority Claimed

☒ ☐  
Yes No

PP3634	Australia	21/05/1998
(Number)	(Country)	(Day/Month/Year Filed)

☒ Yes    ☐ No

(Number)	(Country)	(Day/Month/Year Filed)
1	USA	10/10/2017
2	USA	10/10/2017
3	USA	10/10/2017
4	USA	10/10/2017
5	USA	10/10/2017
6	USA	10/10/2017
7	USA	10/10/2017
8	USA	10/10/2017
9	USA	10/10/2017
10	USA	10/10/2017
11	USA	10/10/2017
12	USA	10/10/2017
13	USA	10/10/2017
14	USA	10/10/2017
15	USA	10/10/2017
16	USA	10/10/2017
17	USA	10/10/2017
18	USA	10/10/2017
19	USA	10/10/2017
20	USA	10/10/2017
21	USA	10/10/2017
22	USA	10/10/2017
23	USA	10/10/2017
24	USA	10/10/2017
25	USA	10/10/2017
26	USA	10/10/2017
27	USA	10/10/2017
28	USA	10/10/2017
29	USA	10/10/2017
30	USA	10/10/2017
31	USA	10/10/2017
32	USA	10/10/2017
33	USA	10/10/2017
34	USA	10/10/2017
35	USA	10/10/2017
36	USA	10/10/2017
37	USA	10/10/2017
38	USA	10/10/2017
39	USA	10/10/2017
40	USA	10/10/2017
41	USA	10/10/2017
42	USA	10/10/2017
43	USA	10/10/2017
44	USA	10/10/2017
45	USA	10/10/2017
46	USA	10/10/2017
47	USA	10/10/2017
48	USA	10/10/2017
49	USA	10/10/2017
50	USA	10/10/2017
51	USA	10/10/2017
52	USA	10/10/2017
53	USA	10/10/2017
54	USA	10/10/2017
55	USA	10/10/2017
56	USA	10/10/2017
57	USA	10/10/2017
58	USA	10/10/2017
59	USA	10/10/2017
60	USA	10/10/2017
61	USA	10/10/2017
62	USA	10/10/2017
63	USA	10/10/2017
64	USA	10/10/2017
65	USA	10/10/2017
66	USA	10/10/2017
67	USA	10/10/2017
68	USA	10/10/2017
69	USA	10/10/2017
70	USA	10/10/2017
71	USA	10/10/2017
72	USA	10/10/2017
73	USA	10/10/2017
74	USA	10/10/2017
75	USA	10/10/2017
76	USA	10/10/2017
77	USA	10/10/2017
78	USA	10/10/2017
79	USA	10/10/2017
80	USA	10/10/2017
81	USA	10/10/2017
82	USA	10/10/2017
83	USA	10/10/2017
84	USA	10/10/2017
85	USA	10/10/2017
86	USA	10/10/2017
87	USA	10/10/2017
88	USA	10/10/2017
89	USA	10/10/2017
90	USA	10/10/2017
91	USA	10/10/2017
92	USA	10/10/2017
93	USA	10/10/2017
94	USA	10/10/2017
95	USA	10/10/2017
96	USA	10/10/2017
97	USA	10/10/2017
98	USA	10/10/2017
99	USA	10/10/2017
100	USA	10/10/2017

☐ Yes    ☐ No

That I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, §1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

United States Application(s)

(Application Serial No.)	(Filing Date)	(Status)-(Patented, pending, abandoned)

(Application Serial No.)	(Filing Date)	(Status)-(Patented, pending, abandoned)

(Application Serial No.)	(Filing Date)	(Status)-(Patented, pending, abandoned)

That all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

I hereby appoint the following attorneys, with full power of substitution and revocation, to prosecute this application and to transact all business in the United States Patent and Trademark Office connected therewith and request that all correspondence and telephone calls in respect to this application be directed to: WELSH & KATZ, LTD., 120 South Riverside Plaza, 22nd Floor, Chicago, Illinois 60606-3913, Telephone No.: (312) 655-1500.

Attorney

Registration No.

A. Sidney Katz  
Richard L. Wood  
Jerold B. Schnayer  
Eric C. Cohen  
Joseph R. Marcus  
Gerald S. Schur  
Gerald T. Shekleton  
James A. Scheer  
Daniel R. Cherry  
Kathleen A. Rheintgen  
Edward P. Gamson  
Thomas W. Tolpin  
Shannon L. Nebolsky  
Mitchell J. Weinstein  
Charles R. Krikorian

24,003  
22,839  
28,903  
27,429  
25,060  
22,053  
27,466  
29,434  
29,054  
34,044  
29,381  
26,700  
41,217  
37,963  
40,687

Full name of sole or one  
joint inventor:

Peter Richard Reeves

Inventor's signature:

P. Reeves

Date:

15/3/2001

Residence and Post Office Address:

20 Mansfield Street

Glebe NSW 2037

AUSTRALIA

Citizenship:

Great Britain

Full name of additional joint  
inventor, if any:

Lei Wang

Inventor's signature:

Lei Wang

Date:

15/3/2001

Residence and Post Office Address:

8A Holt Street

North Ryde NSW 2113

AUSTRALIA



Citizenship:

Australia

Address for Correspondence:

**WELSH & KATZ, LTD.**  
120 South Riverside Plaza  
22nd Floor  
Chicago, Illinois 60606-3913

PCT09

SV

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/701,132A

DATE: 10/23/2001  
 TIME: 10:29:44

Input Set : A:\P30384.app  
 Output Set : N:\CRF3\10232001\I701132A.raw

3 <110> APPLICANT: THE UNIVERSITY OF SYDNEY  
 5 <120> TITLE OF INVENTION: ANTIGENS AND THEIR DETECTION  
 7 <130> FILE REFERENCE: REEVES  
 C--> 9 <140> CURRENT APPLICATION NUMBER: US/09/701,132A  
 C--> 10 <141> CURRENT FILING DATE: 2001-06-15  
 12 <160> NUMBER OF SEQ ID NOS: 68  
 14 <170> SOFTWARE: PatentIn Ver. 2.0  
 16 <210> SEQ ID NO: 1  
 17 <211> LENGTH: 1773  
 18 <212> TYPE: DNA  
 19 <213> ORGANISM: Escherichia coli  
 21 <400> SEQUENCE: 1  
 22 atgcgacgta tagaacgaat accgggggta tggcgtaag cggggcaaaag tttagcattt 60  
 23 attttttggc ttaatgcac gaacagcaac gaggaagggg agtatttcga ccgctagaaa 120  
 24 aaaattctaa aggttgtgag tgacagagac ataacagggt tgacggcgac gaagccgaag 180  
 25 ggtggaagcc caactactaa accgtagact tgaacaacag aaaaatgaat atggcacaag 240  
 26 tcattaatca caacagcctc togtgatca ctcaaaataa tatcaacaag aaccagctcg 300  
 27 cgctgtogac ttctatcgag cgcctctctt ctggtctcgc cattacaagc gctaaagatg 360  
 28 acgctgcggg coaagcgatt gctaaccgct tcacttctaa catcaaaagt ctgactcagg 420  
 29 ccgcacgtaa cgccaacgac ggtatttctc tggcgacgac cactgaaggc gcactgtctg 480  
 30 aatcaacaaa caacttgcag cgtgttctgt aactgaccgt tcaggccaat accggtacta 540  
 31 actctgattc tgacctgtct tcaatacagg acgaaatcaa atcccgctc gatgaaattg 600  
 32 accggtatc ccgtcagact cagttcaacg cggttaattgt tctttccaaa gatggttcaa 660  
 33 tgaaaattca ggttggtgag aatgatggct aaactatctc catcgatctg aagaaaattg 720  
 34 attcttcaac ttgggggctg aatgcttctc cagtttctaa aaactctctt aatgtcagca 780  
 35 atgctatacc atctatcccg caagccgcta gcaatgaacc tgttgatgtt aactcgggtg 840  
 36 atactgatga gtctgcagca atcgacgcca aattgggggt ttccgatacg tcaagcctgt 900  
 37 cgctgcacaa catccttgat aaagatggta aggcaacagc tgattatgtt gtccagtcag 960  
 38 gtaaagactt ctatgctgct tctgttaatg ccgcttcagg taaagtaacc ttaaacacca 1020  
 39 ttgatgttat ttatgatgat tatgcgaacg gtgttgacga tgccaagcaac acaggtcagg 1080  
 40 tgatcaaaag ttcagcagat aaagacggcg cagctcaagg ttttgcaca ctccaaggca 1140  
 41 aaaaactattc tgctgggtgat gcggcgacga ttcttaagaa tggagcaaca gctcttaagt 1200  
 42 taactgtact gaatttaagt gatgttaact atactaatgg taaggtaacc acaactgcga 1260  
 43 ctgagcaatt tgaaggtgct tcaactgagg atccgctggc gcttctggat aaagctattg 1320  
 44 catcagtcga caaattccgg tcttctctag gtgccgtgca gaacgctctc gattccgcta 1380  
 45 tcaccaacct gaacaacacc accaccaacc tgtctgaagc gcagtcocgt attcaggagc 1440  
 46 ccgactatgc gaccgaagtg tcaaacatgt cgaagagcga gatctccag caggcaggtg 1500  
 47 acgcctgctc gctcaaaagc aaccaggtac ccgacgaagt tctgtcaact ttacaaggct 1560  
 48 aatggcctta accgtgctga ccccgccacc gggcggggtt ttctgtctcg caatttaaccg 1620  
 49 ataaccccca aataaccctc catttcaacc actaatgctc gattataaaa ccttcagaaa 1680  
 50 acggataatc atgcgataa ctcatataac gcagggtctg ttatcgtgaa ttactctat 1740  
 51 accgctgaag gtgtaatgga taaacactcg ctg 1773  
 53 <210> SEQ ID NO: 2  
 54 <211> LENGTH: 500  
 55 <212> TYPE: DNA  
 56 <213> ORGANISM: Escherichia coli  
 58 <400> SEQUENCE: 2

ENTERED

## RAW SEQUENCE LISTING

DATE: 10/23/2001

PATENT APPLICATION: US/09/701,132A

TIME: 10:29:44

Input Set : A:\P30384.app

Output Set: N:\CRF3\10232001\I701132A.raw

```

59 aacagcctct cgctgatcac tcagaacaac atcaacaaaa accagtcttc aatgtctact 60
60 gccattgagc gtctgtcttc cggctcgcgt atcaacagcg caaaagatga cgtcgtggcg 120
61 caggcgattg ccaaccgctt cacctctaac atcaaaagtc tgactcaggc agctcgtaac 180
62 gccaacgacg gtatctccgt tgcacagacc actgaaggcg cactgtctga aatcaacaac 240
63 aacctggcag gtatccgtga gctgactgtt cagtcttcta cgggtactaa ctctgaatcc 300
64 gatctgaact caatccagga cgaatataaa tcccgctcgg acgaatttga ccgcgtatcc 360
65 ggctcagacc agttcaacgg cgtgaacgtg ctggcaaaag accgctccat gaaaattcag 420
66 gttggcgcg aogatgggtg aaccatcaac atgcacctga aaaaaattga ctcttctact 480
67 ttaaacctga ctgggtttaa          500
69 <210> SEQ ID NO: 3
70 <211> LENGTH: 500
71 <212> TYPE: DNA
72 <213> ORGANISM: Escherichia coli
74 <400> SEQUENCE: 3
75 ctacagtatgc tgtcacccgc agtacagggt cggtaactta cgatccagat acagatccgt 60
76 ccgcgactgg tgatatgtgt tctgcttatg ttgatgatgc aggtacattg acaactgatg 120
77 caaacaatac tgtaaatatg tatgcccaac ctaatggtag cgtcacgaac gacagtgggt 180
78 cagctattta cgaactgaa cggggcaaat tgactactga agcgtctaca gctcgtgaaa 240
79 ctaccgctaa cccactgaaa gccctggacg atgcaatcag ccagatcgac aaattccgtt 300
80 ctctctgggg tgctgtacag aacctgtcgt attctcgggt aaccaactcg aacaacaaca 360
81 ccaccaacct gtctgaagcg cagtcccgta ttcaggacgc cgaactatgc accgaagtgt 420
82 caaatatgtc taaagcgacg atcatccagc aggcgggtaa ctccgtgttg gctaaagcta 480
83 accaggttcc tcagcaggtt          500
85 <210> SEQ ID NO: 4
86 <211> LENGTH: 399
87 <212> TYPE: DNA
88 <213> ORGANISM: Escherichia coli
90 <400> SEQUENCE: 4
91 agcctgtgcg tgttgacca gaataacctg aacaaatctc agtcttctct gagctccgcc 60
92 attgagcgct tctctcttgg cctgcgtatt aacagtgcta aagatgacgc agcaggtcag 120
93 gcgattgcta accgtttttc agcaaatatt aaaggcttga ctcaggcttc ccgtaacgcg 180
94 aatgatggta tttctgttgc gcagaccact gaaggcgcgc tgaatgaaat taacaacaac 240
95 ctgcagcgct tacgtgaact gactgttcag gcaactaacg gtactaactc tgacagcgat 300
96 ctcttctcta tccagctga aattactcaa cgtctggaag aaattgaccg tgtatctgag 360
97 caaactcagt ttaaccgcgt gaaagtcctt gctgaaaa          399
99 <210> SEQ ID NO: 5
100 <211> LENGTH: 417
101 <212> TYPE: DNA
102 <213> ORGANISM: Escherichia coli
104 <400> SEQUENCE: 5
105 gcacgttagt tgttaacggt gcaacttacg atgttagtgc agatggtaaa acgataacgg 60
106 agactgtctc ttgttaacaat aaagtcatgt atctgagcaa atcagaaggt ggtagcccga 120
107 ttctggtaaa cgaagatgca gcaaatcgtg tgcaatctac caccaaccgc ctcgaaacta 180
108 tcgacaaaagc attggctaaa gttgacaatc tgcgttctga cctcgtgtga ctcacaaaac 240
109 gtttcgactc tgctatcacc aaccttgga caccgctaaa caactgtctc gtgcgccgtg 300
110 gcgattctga agatgctgac tacgcgacgc aagtgctcaa cgtctcgtcgt gcgcagatcc 360
111 tgcaacaagc gggatctctct gttctggcgc aggcatacca gaccacgcag accgtatc 417
113 <210> SEQ ID NO: 6
114 <211> LENGTH: 950

```

## RAW SEQUENCE LISTING

DATE: 10/23/2001

PATENT APPLICATION: US/09/701,132A

TIME: 10:29:44

Input Set : A:\P30384.app

Output Set : N:\CRF3\10232001\I701132A.raw

```

115 <212> TYPE: DNA
116 <213> ORGANISM: Escherichia coli
118 <400> SEQUENCE: 6
119 aacaaaaaac agtctgcgct gtgcacttct atcgagcgcc tctctttctg tctgcgtatt 60
120 aacagcgcta aagatgacgc cgcgggccag gcgattgcta accgctttac ttctaacatc 120
121 aaaggtctga ctacggccgc acgtaacgcc aacgacgcta tttctctggc gcagacggct 180
122 gaaggcgcgc tgtcagagat taacaacaac ttgcagcgta ttcgtaact gaccgttcag 240
123 gctctaccgc gcacgaactc tgattccgac ctgtcttcta ttcaggacga aatcaaatcc 300
124 cgtcttgatg aaattgaccg tgtatctggg cagaccagct tcaacgggtg gaacgtgctg 360
125 tcgaaaaaac attogatgaa gattcagatt ggtgccaatg ataaccagac gatcagcatt 420
126 ggcttgcaac aaatcgacag taccactttg aatctgaag gatttaccgt gtcggcgatg 480
127 gcggaattca gcgcggcgaa actgaaggct gctgatggta cagcaattgc tgcgtcgcat 540
128 gtcaaggatg ctgggggtaa acaagtcaat ttactgtctt acactgacac cgcgtctaac 600
129 agtactaaat atgcggtcgt tgattctgca accggtaaat acatggaagc cactgtagtc 660
130 attaccggta cgcgcggcgc ggtaactgtt ggtgcagcgg aagtggcggg agccgctaca 720
131 gcgcgtccgt taaaagcact ggatgccgca atcgctaaag tcgacaaatt ccgctctccc 780
132 ctccggtgcg ttcaaaaaccg tctggattct gcggtaccca acctgaacaa cctccaccac 840
133 aacctctctg aagcgcagtc ccgtattcag gacgcccact atgcgcagca agtgtccaac 900
134 atgtcgaaag cgcagattat ccagcaggcg ggcaactccg tgcgtcttaa 950
136 <210> SEQ ID NO: 7
137 <211> LENGTH: 1212
138 <212> TYPE: DNA
139 <213> ORGANISM: Escherichia coli
141 <400> SEQUENCE: 7
142 aacaaaaaac agtctgcgct gtgcacttct atcgagcgcc tctctttctg tctgcgtatt 60
143 aacagcgcta aagatgacgc cgcgggccag gcgattgcta accgcttccac ttctaacatc 120
144 aaaggtctga ctacggccgc acgtaacgcc aacgacgcta tctctctggc gcagaccact 180
145 gaaggcgcgc tgtctgaaat caacaacaac ttgcagcggt tgcgtgagtt gaccgttcag 240
146 gcgacgaccg ggactaacct tgattctgac ctgtcttcta ttcaggacga aatcaaatcc 300
147 cgtctggatg aaattgatgc cgtttccggg cagaccagct tcaacggcgt gaattgtcgt 360
148 gcgaaagatg gttcgtgtaa gattcaggtt ggcgcgaaat atgggcagac tttagcatt 420
149 gatttgcaga agattgactc ttctacatta ggactgaacg gtttctccgt ttccgggtcag 480
150 tcaactaaag ttatgtatcc cactactcaa attaccgggt ccgcccggag aaaaactggt 540
151 ggtgttgatt tcaactgctg tgcgaaagat ctgactactg gcagacggtaa aacagtcgat 600
152 gtttctcagc tgacgtttaca caacactctg gatgcgaag gggctgctac atcacagttc 660
153 gtctgtccat ccggcaatga ttctactccc gctgcgatta atcataacga cggcaaatgc 720
154 acgttgaata aagccgatgt cgaataacaca gacaccgata atggactaac gactcgcggt 780
155 actcagaaga atcaactgat taaagttgcc gctgactctg accgctcgcg tcgcggtgat 840
156 gtaacattcc aaggtaaaaa ctacgctaca acggtttcaa cggcacttga tgataaact 900
157 gcggcaaaag caacagataa taaagttggt gttgaattat caacagcaaa accgactgca 960
158 cagttctcag gggctctctc tgcgtatcca ctggcacttt tagacaaagc tattgcacg 1020
159 gttgatactt tccgctctcc cctcgggtgcg gtgcaaaacc ctgtgattc cgagtaaac 1080
160 aacctgaaca acacaccac caactgtctc gaagcgcagt cccgtgattc ggaacgcgac 1140
161 tatgtcacag aagtggtcca catgtcgaaa gcgcagatca tccagcaggc aggttaactg 1200
162 gtgctgtcca aa
164 <210> SEQ ID NO: 8
165 <211> LENGTH: 1647
166 <212> TYPE: DNA
167 <213> ORGANISM: Escherichia coli

```

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/701,132A

DATE: 10/23/2001  
TIME: 10:29:44

Input Set : A:\P30384.app

Output Set: N:\CRF3\10232001\I701132A.raw

```

169 <400> SEQUENCE: 8
170 atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaataa tatcaacaag 60
171 aaccagctcg cgtctgtcag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
172 gcgaaggatg acgcgcgcgg tcaggcgatt gctaaccgtt ttactttcaa cattaaaggc 180
173 ctgactcagg ctgcacgttaa cgccaaacgac ggtattttctg ttgcgcagac caccgaagcg 240
174 gcgctgtccg aaattaacaa caacttacag cgtatttcgt aactgacggt tcaggctctt 300
175 accgggaacta actctgattc ggaatcggac tccattcagg acgaaataca atccgctctc 360
176 gacgaaattg accgcgtatc cggtcagacc cagttcaacg cgggtgaacgt actggcaaaa 420
177 gacggttcga tgaaaaattca ggttggtgcg aatgacggcc agactatcac tattgatctg 480
178 aagaaaaattg actctgatac gctggggctg aatgggttta atgtgaaacg caaaggggaa 540
179 acggctaata cggcagcaac cctgaaagat atgtctggat tcacagctcg ggcgccacca 600
180 ggggggaactg ttggtgtaac tcaatactac gacaaatcgg ctgtgacaag tagcgtagat 660
181 attctaattg ctgttctcgg ccagatgga aataaagtta caactagcgc cgatgttgtt 720
182 tttgtgtacac cagcgcgtgc tgtaacctat acctacaata aagacataca ttcatattcc 780
183 ccgcctctctg atgatatttc cagcgctaac ctggctgctt tctctaatcc tcaggccgga 840
184 gatacgacta aagctacagt tacaattggt ggcaaaagtc aagatgtaaa tccgataaaa 900
185 tccggtaatt taactctgct tgatgatggc gcagtaacct atattgtacg catcgctaac 960
186 ttaactaaaa ataactctgg ttgtgataca caagctactt tggctaaact tgactctgct 1020
187 attcgtgacta aagcgcgcac catccaactc gataaaggaa catccaccag tgcgtaccga 1080
188 cgttttgatg gtgcatacat gtccattgat accaatacat ttgcaaatgc agtaaaaaat 1140
189 gacacttata ctgcacagt aggtgctaac acttatagcg taacacaacg ttctgctgct 1200
190 gcagacacgc cttatatgct caatgggggt ctcagtgata ctcgcgcaac ttactatgca 1260
191 caagctgatg gaagtatac aactactgag gatcgcgctg ccggtaaact ggtctacaaa 1320
192 ggttcgatg gtaagttaac aacggatacg actagcaaac cagaatacac atcagatccg 1380
193 ctggcagctc ttgacgacgc tatcagccag atcgacaaat tccgctctcc cctgggtgcg 1440
194 gtgcaaaaacc gtctggattc cgcagtgacc aacctgaaca acaccactc caactctctc 1500
195 gaagcgcagt cccgtattca ggaacgcgac tatgcgacc aagtgtccaa catgtcgaaa 1560
196 gcgcagatta tccagcagcg cggtaacctc gtgctggcaa aagctaaacca ggttcgcgag 1620
197 caggttctgt ctcgtctgca ggggttaa 1647
199 <210> SEQ ID NO: 9
200 <211> LENGTH: 1758
201 <212> TYPE: DNA
202 <213> ORGANISM: Escherichia coli
204 <400> SEQUENCE: 9
205 atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaataa tatcaacaag 60
206 aaccagctcg cgtctgtcag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
207 gcgaaggatg acgcgcgcgg tcaggcgatt gctaaccgtt ttactttcaa cattaaaggc 180
208 ctgactcagg ctgcacgttaa cgccaaacgac ggtattttctg ttgcgcagac caccgaagcg 240
209 gcgctgtctg aaatcaacaa caacttacag cgtatccgtg agctgacggt tcaggctctt 300
210 accgggaacta actctgattc ggaatcggac tccattcagg acgaaataca atccgctctc 360
211 gatgaaattg accgcgtatc cggcagacac cagttcaacg cgggtgaacgt actggcaaaa 420
212 gacggttcga tgaaaaattca ggttggtgcg aatgacgggt aaactatcac tatcgacctg 480
213 aagaaaaatcg actctgatac tctgggtctg aatgttttta acgtaaatcg taaagattct 540
214 attaccaaca aagctgcaac ggttaagtgt ttaacttctg ctggcgcgaa gttaaacacc 600
215 acgacaggtc tttatgatct gaaaaccgaa aataaccttg taactaccg tgctgacct 660
216 gataaaattg ggaatgcgca taaagtcaac gttggcgcg tagattatac actggcaacg 720
217 aaatctgttg attttactac caccaaatct actgctggtg cgggtgtaga cgccgcgcg 780
218 caggtactcg attcaagctaa aaaaactgta cggttagctg ccaccttca tgcgtatgtg 840
219 ggttaactctg ttaattggtt tataccaca aaagatggtt ctgtttcttt cgaaccgagt 900

```

## RAW SEQUENCE LISTING

DATE: 10/23/2001

PATENT APPLICATION: US/09/701,132A

TIME: 10:29:44

Input Set : A:\P30384.app

Output Set: N:\CRF3\10232001\I701132A.raw

```

220 tcacgacggtta atatcacccat cggttggaagc caggcatatcg tagacgatgc aggcaacttg 960
221 acgactaaca acgctggttag cgcagctataa gctgatatga aagcgctgct taaagccgcg 1020
222 agcgaaggtga gtgacgggtgc ttctctgaca ttcaatggca ctgaataatac tatcgcaaaa 1080
223 gcaactctgc cgacaacctc tccagtagct ccgttaatcc ctggtgggat taacttatcag 1140
224 gctacagtgta gtaaaagtgt agtattgagc gaaaccaaag cggctgccgc gacatcttca 1200
225 attaccttta attccgggtgt actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
226 gatgctgcga agtcttatgt ggatgataaa ggtgggtatta ctaacgttgc cgactataca 1320
227 gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgcccggga tgcttcagcg 1380
228 actgatacca ataaagatta tgctccagca attggtactg ctgtaaatgt gaactccgcg 1440
229 ggtaaaatca ctactgagac taccagtgtc ggttctgcaa cgaccaaccc gcttgcgtgc 1500
230 ctggagcagc cttacagctc catcgacaaa ttccgttctt ccctgggtgc tatccagaa 1560
231 cgtctggatt ccgcagtgac caacctgaac aacaccaata ccaacctgtc tgaagcgagc 1620
232 tcccgtatcc aggacgcgca ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatt 1680
233 atccagcagc ccggtaactc cgtgctggca aaagccaacc aggtaccgca gcaggtttctg 1740
234 ttctctgctgc aggggttaa 1758
236 <210> SEQ ID NO: 10
237 <211> LENGTH: 1383
238 <212> TYPE: DNA
239 <213> ORGANISM: Escherichia coli
241 <400> SEQUENCE: 10
242 aacaaatctc agtctctctc tagctctgct attgagcgtc tgtcttttgg tctgcgtatt 60
243 aacagcgcaa aagacgatgc agcaggctcg gcgattgcta accgttttac ggcaaatatt 120
244 aaaggtctga cccaggtctc ccgtaacgca aatgatggta tttctgttgc gcagaccatt 180
245 gaaggtgcgc tgaatgaaat taacaaacac ctgcagcgta ttctgttaact ttctgttcag 240
246 gcaactaacg tgaactaact tgacagcgat cttttcttta tccaggctga aattactcaa 300
247 cgtctggaag aaattgaccg tgtatctgag caaactcagt ttaacggcgt gaaagtcttt 360
248 gctgaaaata atgaaatgaa aattcaggtt ggtgctaatt atggtgaaac catcaactac 420
249 aatctggcaa aaattgatgc gaaaactctc ggcttgagc gttttaatat cgatggcgcg 480
250 cagaaagcaa caggcgatga cctgatttct aaatttaaag cgacagggtac tgataattat 540
251 gatgttgccg gtaaaactta taccgtgaat gtggagagcg gcgcggttaa gaatgatgct 600
252 aataaagatg tttttgtaag cgcagctgat ggatcgctga cgaccagtag tgatactaaa 660
253 gtatccgggtg aaagtattga tgcaaacgaa ctgacgaaac ttgcaataaa attagctgac 720
254 aaaggctcca ttgataacaa gggcattaca ttactaaca acactggcgc agagcttgat 780
255 gctaattgta aaggtgtttt gaccgcaaat attgatggtc aagatggtca atttactatt 840
256 gacagtaaat caccacaggg tgccggcgca acaataacta cacacacagc tgtttacaaa 900
257 aacagtgccg gcagagtctac cactacaaaa gtggaaaaata aagccgcaac actctctgag 960
258 ctggatctta atgcagcgcaa gaaaacaggt agcactttag ttgtaaatgt cgccacctac 1020
259 aatgtcagcg cagatggtaa aacggttaact gatactactc ctggtgcocc tgaagtgcag 1080
260 tatctgagca aatcagaagg ttgtagcccg attctggtaa acgaagatgc agcaaaatcg 1140
261 ttgcaattca ccaccaaccc gctcgaaact atcgacaagg cattgtgctaa agttgacaa 1200
262 ctgctgtctc acctcggtgc agtacaacaa cgtttctgact ctgccatcac caactctggc 1260
263 aacaccgtta acaacctgtc ttctgcccggt agccgtatcg aagatgctga ctacgcgacc 1320
264 gaagtgtcta acatgtctc tgccgagatc ctgcaacaag cgggtgacctc tgttctggcg 1380
265 cag
267 <210> SEQ ID NO: 11
268 <211> LENGTH: 2013
269 <212> TYPE: DNA
270 <213> ORGANISM: Escherichia coli
272 <400> SEQUENCE: 11

```

**VERIFICATION SUMMARY**

DATE: 10/23/2001

PATENT APPLICATION: US/09/701,132A

TIME: 10:29:47

Input Set : A:\P30384.app

Output Set: N:\CRF3\10232001\I701132A.raw

L:9 M:270 C: Current Application Number differs, Replaced Application Number

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date

{01}  
 {02}  
 {03}  
 {04}  
 {05}  
 {06}  
 {07}  
 {08}  
 {09}  
 {10}  
 {11}  
 {12}  
 {13}  
 {14}  
 {15}  
 {16}  
 {17}  
 {18}  
 {19}  
 {20}  
 {21}  
 {22}  
 {23}  
 {24}  
 {25}  
 {26}  
 {27}  
 {28}  
 {29}

PCT09

RAW SEQUENCE LISTING  
 PRT:PLI APPLICATION US/09/701,132  
 DATE: 01/23/2001  
 TIME: 11.14 28

Input Set A:\P30384.app  
 Output Set M:\CRF3\01233001\1701132.raw

ENTERED

5000

1 <110> APPLICANT THE UNIVERSITY OF SYDNEY  
 5 <120> TITLE OF INVENTION ANTICENS AND THEIR DETECTION  
 7 <130> TITLE REFERENCE NUMBERS  
 C--> 9 <140> CURRENT APPLICATION NUMBER: US/09/701,132  
 C--> 10 <141> CURRENT FILING DATE: 2000-11-21  
 12 <160> NUMBER OF SEQ ID NOS 68  
 14 <170> SOFTWARE PatentIn Ver 2.0  
 16 <210> SEQ ID NO 1  
 17 <211> LENGTH 1773  
 18 <212> TYPE: DNA  
 19 <213> ORGANISM Escherichia coli  
 21 <400> SEQUENCE: 1  
 22 atgcacacga tagaacgaat acccggatla tggcgtaaa cggggcdaaa ttacgattt 60  
 23 attttttagc ttaatgacac gaacagcaac gaqggaaggg agtattttag ccgtacaaaa 120  
 24 aaatttttaa aggttgtaga tgacacgaag alaacagggc tgacggcgac aaagcgaaaa 180  
 25 qgggacgaac caactactaa accgtagact tgaacacaga aaatggaac atggcacaac 240  
 26 tcaattttaa caacagcgcg tcaatgata ctaaaatata ta'caacaa agccagctct 306  
 27 ccgttcgaac ttctatcgaa cagctctctt ctgctctggc ca'caacag cc'caaaatg 366  
 28 accgttcagg caaaggatll gttacacgcl reactttaa caaaagagt ctgcttcagg 429  
 29 cccacgctaa cgcacacaga gcatcttcc cagcgctgac rac'gaaggg gcaatgctgc 480  
 30 aaatcaacaa caacttcagc cgtattctga aactaacat taagacacat accgtaacta 540  
 31 acctctgatt tgacctctat tcaatcaggg acataalcaa atcc'gtctc gatacaatcg 600  
 32 accgctatc cggctcgaat cagctcaagc gcatctatc ttcttccaaa atggttcaa 660  
 33 tgaataatla gtttggtggc aatgatggt aaactatctc cagcgtctgc aaaaatattg 720  
 34 attcttcaaa ttggggctgc aatgcttct ccgttttcaa aaactcttt aaatctcagc 780  
 35 atgctatcaa atcttctcgg caaacgccta gcaatcaacc tgg'gcatgt aaatctcagc 840  
 36 atcaatcaaa gtttcgagca at'caacgcaa atgtggggat tccgattag caaaacacat 900  
 37 cgttcgacaa caatctctgc aaatggttca aggcacacac tgaattglt gttcacttaa 960  
 38 gtaagacat c'atgctctc tctgttata ccgttccggt taagttaacc taaacaccaa 1020  
 39 gtaagacat c'atgctctc tctgttata ccgttccggt taagttaacc taaacaccaa 1080  
 40 tgaatgacat tgaatgacat aaatcaggg ccgttcaggg ttgttccaa cttcagagga 1140  
 41 aaactatc tgggttgtag cggcgagaca ttcttcagaa tgaagacaa acctttaaag 1200  
 42 taacttactt gaatttaagt gattgtctgc atactaatgg taagttaacc aaatataaa 1260  
 43 ctgagcaatt taagtattgt tcaactcagg accgctcggc actcttgat aaagtattg 1320  
 44 caatattcga aaatttcagg tcttctctag atgccttcaa gaacgtctc gattccctta 1380  
 45 tcaacacatc gaagacacac accacacac tctctgaag gaaagccgtt attcagcagc 1440  
 46 caactatcga gacgcaatt tcaacacatc gaaagacga gactatcag caggatgaa 1500  
 47 atctctgct gttcaagc aaacagac cgaacacat tctgacatg ttcaatgct 1560  
 48 atagctttaa acatgcttaa cccacacac cggcgatll ttctgctcc caattttaa 1620  
 49 ataaacccca aaataacacat catcttcaa actatctgc gattataaa ccctcagaa 1680  
 50 accgataatc atgcgataac c'catctaac gaaagctgt ttatcttcaa tcaactttaa 1740  
 51 accgttcagg gtttcaagga taacacatc ctg 1773  
 53 <210> SEQ ID NO 2  
 54 <211> LENGTH 506  
 55 <212> TYPE: DNA  
 56 <213> ORGANISM Escherichia coli  
 58 <400> SEQUENCE: 2



## DATE 02/24/2001

JUNE 12 14 28

Output Set: N:\CRF3\01232001\I701132.raw

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)	(46)	(47)	(48)	(49)	(50)	(51)	(52)	(53)	(54)	(55)	(56)	(57)	(58)	(59)	(60)	(61)	(62)	(63)	(64)	(65)	(66)	(67)	(68)	(69)	(70)	(71)	(72)	(73)	(74)	(75)	(76)	(77)	(78)	(79)	(80)	(81)	(82)	(83)	(84)	(85)	(86)	(87)	(88)	(89)	(90)	(91)	(92)	(93)	(94)	(95)	(96)	(97)	(98)	(99)	(100)
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	

RAW SEQUENCE LISTING  
 PATENT APPLICATION US/09/701,132  
 DATE 01/23/2001  
 TIME 12:14:25

Input Set A:\P30384.app  
 Output Set N:\CRF3\01232001\I701132.raw

```

115 <212> NYPB. DNA
116 <213> ORGANISM Escherichia coli
117 <400> SEQUENCE: 5
118 aacaaacacg agtctacgct gtacagctct atagagcgcg tctctctctg tctgctgatt 60
119 aacagcgcta agatagacgc cggcgacgag gcatgttgta acgctctaac ttctaacatc 120
120 aaggtctgta ctacagcgcc accgaacgac aacgaagcta ttctctctcg gcagacgact 130
121 aacggcgccg tctcaaaatc taccacacac ttacagcta ttctggaact gaacgtttac 140
122 acgcttaacg acacaaacgc caattctaac ctactctcta tttaacgata aataaacatc 150
123 caattctgag aattgagcgg tctatctggt cagacacgac tttaacggtt gaacgtgtcg 160
124 tcaaaacgct atagatgata natctacgct gctgcgaatg ataacgcgac gatacgaatt 170
125 aggtctaacg aatctgacag caatctctac aattctaacg gactctacct gtctgcgata 180
126 aacgtctgta cggcgcgaa atgcaacgct gctactcata cagcaattac tctctccatc 190
127 atcaaatgta ctggcgcgaa acaactctat ttactctctt acgtctaacg cgaactctac 200
128 aattctaaat acgctctctc tcaattctga accggttaat aatctgaagc cactgttagt 210
129 attctacgta cggcgcgacg gctacgctt tttatagcgga aagtgcggcg agcgtctata 220
130 ggcgtctcgt ttacacacat agatctcgca ttcaatlaag tctgcaaaat ccaactctcc 230
131 ctacgttgca tcaaaacgct atggaattct gcatctacac aactgaacaa caccacacac 240
132 aacgtctcga agacgcaatc ccaattctca gacgcgcacg atcgagacga agtctccaac 250
133 aatctgaagc cgaacattat ccaagacgacg ggcacactcg tgcgtctctaa 260
134 <215> STOP ID NO 7
135 <215> LBNM 1212
136 <212> NYPB. DNA
137 <213> ORGANISM Escherichia coli
138 <400> SEQUENCE: 7
139 aacaaacacg agtctacgct gtacagctct atagagcgcg tctctctctg tctgctgatt 60
140 aacagcgcta agatagacgc cggcgacgag gcatgttgta acgctctaac ttctaacatc 120
141 aacgtctgta ctacagcgcc accgaacgac aacgaagcta ttctctctcg gcagacgact 180
142 aacggcgccg tctcaaaatc taccacacac ttacagcta ttctggaact gaacgtttac 240
143 acgcttaacg acacaaacgc caattctaac ctactctcta tttaacgata aataaacatc 300
144 caattctgag aattgagcgg tctatctggt cagacacgac tttaacggtt gaacgtgtcg 360
145 tcaaaacgct atagatgata natctacgct gctgcgaatg ataacgcgac gatacgaatt 420
146 aggtctaacg aatctgacag caatctctac aattctaacg gactctacct gtctgcgata 480
147 aacgtctgta cggcgcgaa atgcaacgct gctactcata cagcaattac tctctccatc 540
148 atcaaatgta ctggcgcgaa acaactctat ttactctctt acgtctaacg cgaactctac 600
149 aattctaaat acgctctctc tcaattctga accggttaat aatctgaagc cactgttagt 660
150 attctacgta cggcgcgacg gctacgctt tttatagcgga aagtgcggcg agcgtctata 720
151 ggcgtctcgt ttacacacat agatctcgca ttcaatlaag tctgcaaaat ccaactctcc 780
152 ctacgttgca tcaaaacgct atggaattct gcatctacac aactgaacaa caccacacac 840
153 aacgtctcga agacgcaatc ccaattctca gacgcgcacg atcgagacga agtctccaac 900
154 aatctgaagc cgaacattat ccaagacgacg ggcacactcg tgcgtctctaa 960
155 <215> STOP ID NO 8
156 <215> LBNM 1614
157 <212> NYPB. DNA
158 <213> ORGANISM Escherichia coli
159 aacaaacacg agtctacgct gtacagctct atagagcgcg tctctctctg tctgctgatt 60
160 aacagcgcta agatagacgc cggcgacgag gcatgttgta acgctctaac ttctaacatc 120
161 aacgtctgta ctacagcgcc accgaacgac aacgaagcta ttctctctcg gcagacgact 180
162 aacggcgccg tctcaaaatc taccacacac ttacagcta ttctggaact gaacgtttac 240
163 acgcttaacg acacaaacgc caattctaac ctactctcta tttaacgata aataaacatc 300
164 caattctgag aattgagcgg tctatctggt cagacacgac tttaacggtt gaacgtgtcg 360
165 tcaaaacgct atagatgata natctacgct gctgcgaatg ataacgcgac gatacgaatt 420
166 aggtctaacg aatctgacag caatctctac aattctaacg gactctacct gtctgcgata 480
167 aacgtctgta cggcgcgaa atgcaacgct gctactcata cagcaattac tctctccatc 540
168 atcaaatgta ctggcgcgaa acaactctat ttactctctt acgtctaacg cgaactctac 600
169 aattctaaat acgctctctc tcaattctga accggttaat aatctgaagc cactgttagt 660
170 attctacgta cggcgcgacg gctacgctt tttatagcgga aagtgcggcg agcgtctata 720
171 ggcgtctcgt ttacacacat agatctcgca ttcaatlaag tctgcaaaat ccaactctcc 780
172 ctacgttgca tcaaaacgct atggaattct gcatctacac aactgaacaa caccacacac 840
173 aacgtctcga agacgcaatc ccaattctca gacgcgcacg atcgagacga agtctccaac 900
174 aatctgaagc cgaacattat ccaagacgacg ggcacactcg tgcgtctctaa 960
175 <215> STOP ID NO 9
176 <215> LBNM 1614
177 <212> NYPB. DNA
178 <213> ORGANISM Escherichia coli

```

```
PAW SEQUENCE LISTING                                DATE: 01/21/2001
PATIENT APPLICATION    US/09/701,132                TIME: 11 34 25

Input Set      A:\P30384.app
Output Set     N:\CRF3\01232001\I701132.raw
```

[illegible]

RAW SEQUENCE LISTING  
 PATENT APPLICATION US/09/701,132

DATE: 01/23/2001  
 TIME: 11.14.29

Input Set: A:\P30384.app  
 Output Set: M:\CRF3\01232001\701132.raw

```

220 tcaagcagga atataacat cagtcgaacc cagcagcaga tagacagaga aagaacatta 959
221 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 1020
222 aagcagagga atataacat cagtcgaacc cagcagcaga tagacagaga aagaacatta 1080
223 gcaactctta caataacat cagtcgaacc cagcagcaga tagacagaga aagaacatta 1140
224 gcaactctta caataacat cagtcgaacc cagcagcaga tagacagaga aagaacatta 1200
225 attacattta attacattta attacattta attacattta attacattta attacattta 1260
226 attacattta attacattta attacattta attacattta attacattta attacattta 1320
227 attacattta attacattta attacattta attacattta attacattta attacattta 1380
228 attacattta attacattta attacattta attacattta attacattta attacattta 1440
229 attacattta attacattta attacattta attacattta attacattta attacattta 1500
230 attacattta attacattta attacattta attacattta attacattta attacattta 1560
231 attacattta attacattta attacattta attacattta attacattta attacattta 1620
232 attacattta attacattta attacattta attacattta attacattta attacattta 1680
233 attacattta attacattta attacattta attacattta attacattta attacattta 1740
234 attacattta attacattta attacattta attacattta attacattta attacattta 1758
235 <10> SEQ ID NO: 10
236 <11> LENGTH: 1383
237 <12> TYPE: DNA
238 <13> ORGANISM: Escherichia coli
239 <14> SEQUENCE: 10
240 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 60
241 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 120
242 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 180
243 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 240
244 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 300
245 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 360
246 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 420
247 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 480
248 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 540
249 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 600
250 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 660
251 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 720
252 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 780
253 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 840
254 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 900
255 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 960
256 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 1020
257 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 1080
258 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 1140
259 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 1200
260 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 1260
261 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 1320
262 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 1380
263 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 1383
264 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 1383
265 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 1383
266 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 1383
267 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 1383
268 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 1383
269 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 1383
270 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 1383
271 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 1383
272 aagcagataga accgtagagc cagcagataga gtagacagaga aagcagataga taacacagag 1383

```

DATE 31/23/2001

Page 12 of 30

Output Set: N:\CRF3\01232001\I701132.raw

[illegible]

## SEQUENCE LISTING

&lt;110&gt; THE UNIVERSITY OF SYDNEY

&lt;120&gt; ANTIGENS AND THEIR DETECTION

&lt;130&gt; REEVES

&lt;140&gt;

&lt;141&gt;

&lt;160&gt; 68

&lt;170&gt; PatentIn Ver. 2.0

&lt;210&gt; 1

&lt;211&gt; 1773

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 1

```

atgcgacgta tagaacgaat accgggggta tcggcgtaag cgggggcaag tttacgatt 60
atcttttggc ttaatgacac gaacagcaac gaggaagggg agtatttcga ccgctagaaa 120
aaaattcttaa aggttgtgag tgaccagacg ataacagggg tgacggcgac gaagccgaag 180
gggggaagcc caatacttaa accgtagact tgaaaaacagg aaaaatgaatc atggcacaag 240
tcatttaaac caacagcctc tcgctgatca ctcaaaaata tatcaacaag aaccagctgt 300
cgctgtcgac ttctatcgag cgctctcttt cgctctctgt cattaacagc gctaaagatg 360
acgctgcggg ccaagcgatt gctaaacgct tcactttcaa catcaaaagg ctgactcagg 420
ccgcacgtaa cgcaacgac tggtattttc tggcgagac cactgaaggc gcactgtctg 480
aaatcaacaa caacttgca gctgttcgtg aactgaccgt tcaggccact accggtacta 540
atcttgattc tgacctgtct tcaatacagg acgaaatcaa atccgctctc gatgaattg 600
accggtatc cggtcaact cagtcaacg cggttaagt tctttccaaa gatggttcaa 660
tgaaaattca gggttggtcg aatgatgggt aactatctc caactttcaa aactctctct aatgtcagca 720
atcttccaac ttgggggctg aatggtctct gactttctaa aactctctct aatgtcagca 780
atgctatcac atctatcccg caagcgcgta gcaatgaacc tggttgatgt aactctgggtg 840
atactgatga gctgcagca atcgacgaca aatgggggtg tcccgatacg tcaagccgtg 900
cgctgcacaa catccttgat aaagattggt ggcgaacagc tgattatgtt gttcagtcag 960
gtaagagact ctatgctgt tctgttaatg ccgcttcagg taaagtaacc ttaaacacca 1020
ttgatgtttc ttatgatgat tatgcgaacg gtgttgacga tgccaagcaa acaggctcagc 1080
tgatcaaaag tgatcgagat aaagacggcg cagctcaagg ttttgtcaca ctccaaggca 1140
aaaaactatc tgcgtggtgat gcggcagaca ttctaaagaa tggagcaaca gctcttaagt 1200
taactgatct gaatttaagt gatgttactg atactaatg taaggtaacc acaactgcga 1260
ctgagcaatt tgaaggtgct tcaactgagg atccgctggc gctctggatg aaagctattg 1320
catcagtcga caaattccgg tcttctctag gtgcgctgca gacccgtctc gatctcgcta 1380
tcaccaacct gaaccaacacc accaccaaac tgcctgaagg gacgctccgt atccaggaga 1440
ccgactatgc gaccgaagtg tccaactgtg caaagcgaga gatctccag caggcaggta 1500
actccgtgct gtctaaagcg aaccaggtac cgacgaaggt tctgtcactg ttaacagggt 1560
aatggcctta acgtgcctga ccccgccacc ggccgggttt ttctgtcgcg caattttacg 1620
ataaccccca aataaccctt catttcacc cactaatcgt cgattaaaaa ccttgacgaa 1680
acggataatc atgccgataa ctcatataac gcagggctgt ttatcgtgaa ttcactctat 1740
accgctgaag gtgtaatgga taaacactcg ctg 1773

```

&lt;210&gt; 2

&lt;211&gt; 500

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 2

```

aacagcctct cgctgatcac tcgaacaac atcaacaaaa accagctctc aatgtctact 60
gccattgagc gctctgtctt ccggtctcgt atcaaacagg caaagatga cgtcgtggc 120
caggcgatgg ccaaccgctt cactcttaac atcaaaagtg tgactcaggc agctcgtaac 180
gccaacgacg gtatctccgt tgcacagacg atcgaaggcg cactgtctga aatcaacaac 240
aacctgcagc gtatccgtga gctgactgtt cagtctctta cgggtactaa ctctgaatcc 300
gatctgaact caatccagga cgaaatataa tccgctctgg acgaaattga cgtcgatccc 360
ggtcagaccc agttcaacg cgtgaacgtg ctggcacaag accgctccat gaaaattcag 420
gttggcgaga acgatggta aaccatcacc atcgacctga aaaaaattga ctctctact 480
ttaaaccctga ctgggtttaa

```

<210> 3  
 <211> 500  
 <212> DNA  
 <213> Escherichia coli

<400> 3  
 ctacagtattgc tgtcaaccgc agtacagggt ccgtaactta cgtaccagat acagatccctg 60  
 cgcgactggc tgataattgt ttgccttatg ttgatgatgc aggtacattg acaactgatg 120  
 caaacaataac tgtaaaatat tatgccaca ctaatggtag cgtccagaa gacagtgggt 180  
 cagctatttta cgcactgaa gcgggcaaat tgactactga agcgtctaca gctgctgaaa 240  
 ctaccgctaa ccactgaaa gcctggagc atgcaatcag ccagatcgac aaattccggt 300  
 ctctctgggg tgcgtatag aacggtctgg attctcggtt aaccaacctg aacaacacca 360  
 ccaccaacct gctcgaagc cagtcocgta ttccaggagc cgtactatgc accgaagtgt 420  
 caaataatgc taaagcgag atcatccagc aggcgggtaa ctccgtgtgt gctaaaagcta 500  
 accaggttcc tcagcaggtt

<210> 4  
 <211> 399  
 <212> DNA  
 <213> Escherichia coli

<400> 4  
 agcctgtgctg tgttgaccca gaataacctg aacaaatctc agtcttctct gagctccgcc 60  
 attgagcgtc tctcttctgg cctgcgtatt aacagtgtcta aagatgagcg accaggttcag 120  
 gcgattgtcta accgtttttc agcaaatatt aaagtgctga ctcaggcttc ccgtaacgcg 180  
 aatgatgtga ttctgttgc gcgagccact gaaggcgccg tgaatgaaat taacaacaa 240  
 ctgcagcggtg tacgtgaact gactgttcag gcaactaacg gtactaaact tgacagcgat 300  
 ctctctctca tcaggctga aattactcaa cgtctggaag aaattgaccg tgtatctgag 360  
 caaactcagt ttaacggcgt gaaagtcctt gctgaaaa

<210> 5  
 <211> 417  
 <212> DNA  
 <213> Escherichia coli

<400> 5  
 gcacgttagt tgttaacggt gcaacttacg atgttagtgc agatggtaaa acgataacgg 60  
 agactgcttc tggtaacaa aaagtcatgt atctgagcaa atcagaagggt ggtagcccg 120  
 ttctgtgtaa cgaagatgca gcaaaatcgt tgcattctac caccaccccg ctgcgaaact 180  
 tcgacaaagg ctgtgctaaa gttgacaatc tgcgttctga cctcgtgtgca gtacaaaaac 240  
 gttctgactc tgctatcacc aaccttgcca acacggtga caactgtct ctgcgccgat 300  
 gccgtatcga agcgtctgac tacgcgaccg aagtgtctaa catgtctcgt gcgcagatcc 360  
 tgcacaaagg gggtacctct gttctggcgc aggtcaacca gaccacgag aacgtac 417

<210> 6  
 <211> 950  
 <212> DNA  
 <213> Escherichia coli

<400> 6  
 aacaaaaacc agtctgcgct gtcgacttct atcagagcgc tctcttctgg tctgcgtatt 60  
 aacagcgcta aagatgacgc cgcggggccag cgtattgcta acgctttac ttctaacatc 120  
 aaaggtctga ctcaggccgc acgtaacgac aacgagcgta ttctctggc gcagacggct 180  
 gaaggcgccg tgtcagagat taacaacaaac ttgcagcgtg ttctgagcga gaccgttcag 240  
 gcctctacgc gcacgaactc tgattccgac ctgtctctta ttcagaga aatcaaatcc 300  
 cgtcttgatg aaattgacgc tgtattctgg cagaccaggt tcaacgggtg gaacgtgctg 360  
 tcgaaaaaac attcgatgaa gattcagatt ggtgccaatg ataacagac gatcagatt 420  
 gccttgcaac aaactgcacg taccactttg aatctgaaag gattacag gtccggcagt 480  
 gcggtattca gcgcggcgaa actgacggct gctgatggta cagcaattgc cgtctcgat 540  
 gtcaaggatg ttgggggtta acaagtcaat ttactgtctt aactgacac cgtctctaac 600  
 agtactaaat atcggtctgt tgattctgca accggtaaat acatggaagc cagtgtatgc 660  
 attaccggta cggcgccggc ggttaactgt ggtgcagcgt atcagcggg agcgcgtaca 720  
 gcgactcgtg taaaagcaat ggatgcgcga atcgtcaaa tcgacaaat cgtctctcc 780  
 ctccgtggcc ttcaaaacgc tctggtattc cgggtcacca accgaacaa caccacacc 840  
 aacctgtctg aagcgcagtc ccgtattcag gacgccgact atgcgacga agtgtccaa 900  
 atgtcgaagg cgcagattat ccagcaggcg ggcactccg tgctgtctaa 950

<210> 7  
 <211> 1212  
 <212> DNA  
 <213> Escherichia coli

<400> 7  
 acaaaaaaac agtctgcgct gtcgacttct atcgagcgcc tctcttctgg tctgcgtatt 60  
 aacagcgcta aagatgaagc gcgaggccag ggcattgcta accgcttcac tcttaacatc 120  
 aaagggtctga ctcaggccgc acgtaaacgc aacgacgctg tctctctggc gcagacccact 180  
 aaaggcgcgcg tgcctgaatc caacaaacac ttgcagcggt tgcgtgaagt gacgcttcag 240  
 gcgacgacgc ggactaacct tgattctgcac cagaccagct ttcaggagca aatcaaatcc 300  
 cgtctggatg aaattgactc cgatttcgggt ggcgcgaagt atgggcagac tattagcatt 420  
 gogaagaatg gttcgatgaa gattcagggt ggcactgaa gtgtctcgt tccgggtcag 480  
 gatttgcaga agattgactc ttctacatta ggaactgaac gtgtctcgt tccgggtcag 480  
 tcaattaacg ttagtgtatc cattaactca attacgggt ccgcggggac aaaactcgtt 540  
 ggtgtgtatt tcaactgctg tgcgaagaat ctgactactg cgacaggtaa aacagtcgat 600  
 gtttctagcc tgacgttaca caacactctg gatgcgaag gggctgctac atcacagttc 660  
 tcogttcaat ccggcaatga ttctactacc gcgtcgatta atcatacaga cggcaaatgc 720  
 acgttgtaata aagccgatgt cgaatacaca gacacggata atggactaac gactgcgctg 780  
 acctcaaaaag atcaactgat taagttgccc gctgactctg acggctcggc tgcgggatat 840  
 gtaacattcc aaggtaaaaa ctacgctaca acggtttcaa cggcacttga tgataaact 900  
 gcggcaaaaag caacagataa taagtttgtt gtgcaaatat caacagcaaa accgactcca 960  
 cagttctcag gggtctcttc tgcgtatcca ctcggcaact tagacaaaagc tattgcacag 1020  
 gtgtgatact tccgctcttc cctcggtggc gtgcaaaaac cctctgattc cgcagtaaac 1080  
 aacctgaaca acaccaccac caacctgtct gaagcgagc cctgat tca ggacggcagc 1140  
 tatgtctacg aaggtgccaa catgtcgaaa gcgcagatca tccagcaggc aggttaactgc 1200  
 gtgctgtcca aa 1212

<210> 8  
 <211> 1647  
 <212> DNA  
 <213> Escherichia coli

<400> 8  
 atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaata tatcaacaag 60  
 aaccagctcg cgcgtctgag ttctatcgag cgtctgtctt ctggcttgcg tattaacacg 120  
 gcgaaggatg taagcgcggg tcaaggcgat gctaaccggt ttaacttcaa cattaaaggc 180  
 ctgactcagg ctgacgtgaa cgcacaagac ggtatttctg ttgcgcagac caaccgaaggc 240  
 cgtctgtcgg aaattataca caacttacag cgtattctgt aactgaacgt tcaggctctc 300  
 accgggacta acctcgatcc ggaatcgagc tccattcagg acgaaatcaa atcccgctc 360  
 gacgaatttg acccgctatc cggtcagacc cagttcaaac cgtgaacgt actggcaaaa 420  
 gacggtctga tgaataatca ggttggtagc aatgacggc agactatcac taatgatctg 480  
 aagaataatg acctcgatac gctggggctg aatgggttta atgtgaacg caaaggggaa 540  
 accgctataa cgcgcagcac cctgaagaat atgtctggat tcaacgctgc ggcggccaaca 600  
 gggggaactg cgtgtgtaac tcaataactc gacaatcgg ctgtgacaag tagcgtagat 660  
 atttctaaag ctgttgcctg cgcagatgga aataaagtta caactacgac cgatgttggc 720  
 ttgtgataca cagcgcgtgc tgtaacctat acctacaata aagacactaa ttcatactcc 780  
 gcgcgtctcg atgatatctc cagcgcctaac ctggctgctt tctcaatcc ttcagctgac 840  
 gatacactaa aagctacagt tacaattggt ggcaaaagct aagatgtaa catcgataaa 900  
 tccgcttaat taactgctgc tgatgatggc gcagtaactt atatggatgc taccggtaac 960  
 ttaactaaaa ataactgctg tggatgata caagctactt tggctaaact tgcactgct 1020  
 actggtgtcta aagcgcgcac catccaactc gataaaggaa cattcaccag tacccgtaca 1080  
 cgcgtttgatg gtgcatacat gtccattgat accaatatcat ttgcaaatgc agtaaaaaat 1140  
 gacacttata ctgcacactg aggtgctaag acttatagcg taacacaagg tctgtgctgc 1200  
 gcgacacagc cttattatgac caatgggggt ctcagtgata ctccgccaac ttactatgca 1260  
 caagctgctg aagtatcac aactactgag gatgcggctg ccggtaaaat cgtctgctgc 1320  
 gggttccgatg tgaagttaac aacggataac actagcaaac cagaatcaac atcagatccg 1380  
 ctggcgctct tgaagcagc tatcagccag atcgacaact tccgctctc atcggttgcg 1440  
 gtgcaaaaac tctggattc cgcagtgacc aacctgaaca acaccactac caacctgtct 1500  
 gaagcgcagc cctgtattca ggacggcagc tatgcgacgc aagtgccaaa catgtcgaaa 1560  
 gcgcagatta tccagcaggc cgttaactcc gtgctgccc aaagtaacca ggttccgacg 1620  
 caggttctgt cctgctgca ggggttaa 1647



<210> 9  
 <211> 1758  
 <212> DNA  
 <213> Escherichia coli

<400> 9  
 atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60  
 aacacgctcg cgctgctcgag ttctatcgag cgctctgctt ctggcttgcg tattaacacg 120  
 gcgaaggatg acgcgcgcgg tcaggcgatt gctaaccggt ttactttcra cattaaaggc 180  
 ctgactctcg cgcgacgtaa cgcacaacgac ggtatttctg ttgcacagac caccgaaggc 240  
 gcgctgctcg actctgattc ggaatcggaac cgtatccgtg agctgaacgt tcaggctctc 300  
 accggaaacta acgcgcgtatc cggccagacc cagttcaaac gcgtgaacgt actgcgcaaa 420  
 gatgaagtgt tgaaaattca ggttgggtgc aatgcagctg aaactatcac tatcgactcg 480  
 aagaaaatcg attctgtatac tctgggtctg aatgggttta acgtaaatcg taaaggctact 540  
 attaccaaca aagctgcaac ggttaagtgt ttaacttctg taactaccga tgcgtcattc 600  
 acgacaggctc ttttatgatc gaaaaccgaa aataccttgt taactaccga tgcgtcattc 660  
 gataaaattag ggaattggcg taaagtcaac actgctggta cgggtgtaga cgccgcgcgc 720  
 aaactctggtg atttactcac caccaaactc actgctggta cgggtgtaga cgccgcgcgc 780  
 caggctactg attcagctaa aaaacgctgac gctgttagct ccacccttca tgcgtgatgt 840  
 ggtaaactcg ttaattggtc ttacaccaca aaagatggta ctgtttcttt cgaaaaggat 900  
 tcacagagta atataccat cgggtggaag caggcatcac tagacgatgc aggcacttg 960  
 acgataacca acgctggtag cgcagctaaa gctgatatac aagcgtgctc taaaggcgcg 1020  
 agcggaagta gtgacgggtc ttctctgaca ttcaatggca ctgaatatac ttacacacc 1080  
 gcaactctcg cgcacaactc tccagtagct ccgttaactc ctgggtggat taactatcac 1140  
 gctacagctga gtaaaagatg agtattgagc gaaaccaaac cggctgcgcg taactatcac 1200  
 attacatttca attccgggtg actaggcaaa actattgggt ttaccgcggg tgaatccagt 1260  
 gatgctcgga agtcttatgc ggaatgataa ggtggtatta ctacgttgc cgactataca 1320  
 gtctctttaca ggtttaacaa ggaataacgc tctgtgactg ttgcgcggta tgcctcagcg 1380  
 actgatacca ataaagatta tgcctcagca attgttactg ctgtaaatgt gaactccgcg 1440  
 ctggacagct ctatcagctc taccagtgct ggttctgcga gcgcaaaccc gcttgcgtcc 1500  
 cgttcggagt ccgcagctac caactgtgaa aacacactca ccaactgttc tgaaggcagc 1560  
 tcccgatctc aggcgcgcga ctatgcgacc gaagtgtcca catgtctcaa agcgcagat 1620  
 attccagcag aggttaactc cgtgctggca aaagccaacc aggtaccgca gcaggttctg 1680  
 tctctgctgc aggggttaa 1740  
 1758

<210> 10  
 <211> 1383  
 <212> DNA  
 <213> Escherichia coli

<400> 10  
 aacaaatctc agtcttctct tagctctgctc attgagcgctc tgtcttctcg tctcgctatt 60  
 aacacgcgca aacagcgatgc agcaggtcag gcgatttgcta accgttttca ggcaaatatt 120  
 aaaggctctga ccaggctctc ccgttaacgca aatgatgtgta ttctctgttc gcagacctact 180  
 gaattgctgcg tgaatgaaat taacaacaac ctgcagcgcta ttctgtgaatc ttctgttcag 240  
 goaactaacg gtaactaact tgacagcgat ctcttctctta tccagctgta tccagctcaa 300  
 cgtctggaag aaattgacgc tgtattctgag caaaactcagt ttaacggcgt gaagctctc 360  
 gctgaaataa attgaaatgaa aattcaggtt ggtgctaatt atggtgaaac catcaactac 420  
 aatctggcga aaattgatgc gaaaactctc ggcttggagc gttttaatat cgatggcgcg 480  
 cagaaaacga caggcagtgta cctgatcttc aaatttaaac cgacaggtga tgataactat 540  
 gatgttgcgc gtaaaactta taccgtgaat gtggagagcg gcgcggttaa gaatgatgct 600  
 aataaagatg ttttggtaag cgcagctgat ggaatcgcta gcacacgatg tgataactaa 660  
 taatccgggtg aaagtattga tgcaacagaa ctacgcgaaac ttgcaataaa attagctgac 720  
 gaaagctcca ttgaatacaa gggcatatac ttactaaca acactggcgc agagcttgat 780  
 gctaatggta aaggtgtttt gaccgcgaat attgatggtc aagatgttca atttactatt 840  
 gcagctaatg caccacoggg tgccgcgcga caataactca cagacacagc ggtttacaaa 900  
 aacagctgcg gccagttcac cactacaaaa gtggaaaaata aagccgcgaac actctctgat 960  
 ctggattctta attgcagcaa gaaaacaggt agcaactttag ttgtaaatgt cggcaactac 1020  
 aatctcagcg cagatggttaa aacggttaact gatactactc ctggtgcccc taaagtgatg 1080  
 tatctgagca aatcagaagg ttgtagcccg attctggtaa acgaagatgc acgaaaatcg 1140  
 ttgcaactta ccacacaccc gctcgaactc atgcgaaagg cattggctaa agttgacaa 1200  
 ctgctgtctg actctcggtgc agtacaacac cgtttctgact ctgcaatcac caactctg 1260  
 aacaccgtaa acaactgttc ttctgccccg agccgctatc aagatgctga ctacgcgacc 1320  
 gaagtgtcta acatgtctcg tgcgcagatc cgtcaacaac cgggtacccc tgttctggcg 1380  
 cag 1383

<210> 11  
 <211> 2013  
 <212> DNA  
 <213> Escherichia coli

<400> 11  
 atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaataa tatcaacaag 60  
 aaccagctctg cgtctgcgag ttctatcgag cgtctgtctt ctggcttgccg tatcaacagc 120  
 gcgaaggatg acgcgcgggg tcaggcgatt gctaaccgtt ttacttctaa cattaaaggc 180  
 ctgactcagg ctgcacgtaa ccccaacgac ggtatttctg ttgcgacagc caccgaaggc 240  
 gcgctgtccg aaattaaacaa caacttacag cgtgtctgtg aactgacggt tcaggctctc 300  
 accgggagcta actccgattc ggaactggac cagttcaacg gcgtgaaact gctgtccaaa 420  
 gacgaatttg accgcgtatc cggccagacc aacgagtgcc aaacgattac tatgtatgtc 480  
 gatggctcga tgaataatca ggtcggcgcc gctgggttta acgttaacgg taaagggtct 540  
 aagaaaattg actctgatac gctgaattct ttaaaactgg ctggtttcac taagggcacc 600  
 gtgacgaata cagctcgacg aagcgacgat aacacaata gtaatgaca agcctaaagt 660  
 acagatacca atggcgtgac cgcgtatata ctggtgatca ctgggggagg ggcaaacgct 720  
 tccgactctg tagctaatat caccgatgga tcagtgatca ctgggggagg ggcaaacgct 780  
 ttggcgtggt ctgcaaaagaa tgggttacaac tatgtatgca caagtaaatc ttatagtctt 840  
 cgtcgagatg gtgcgcatcc agcgaagacg taaagcagca aagttaatat ttcccaggat 900  
 tctgtcgagg cgacagtgac tattgggtgt tagataaaca gggaaactgt 960  
 ggaaaaatca ctgcggcgaga tgataatgag acgctgtatt tagataaaca gggaaactgt 1020  
 accaaaaacga atgcaggtaa cgataccgca gcaactacaa ttcaaatcac ttccggtaac 1080  
 gatcttccag ctggcgtttcc agttgggggt gaaggaaatc agacctcaac aaattctcag 1140  
 gcttccggaa tgctctgttca gtccgcagga gtaagtgat agggagggcg gtctacagac 1200  
 attcttgcag gtggtgcatt tgcgcgtaag gtaagtgat gtatgtcact ttactttgat 1260  
 attttggtag caagttaattg aaacataaca gcggtaaatc caggttccgt ttatgtatgt 1320  
 cgactactcg gtgatttccac tacaacggct caggttaccg ccagaaactc 1380  
 ttaatttgcta acagtaagga tgctacctta acttaccagt ttccacagtt agacacagct 1440  
 gtttatagca caacaggaag tggcgtccag ggtatggcaa atctgacaaa aagcaatttt 1500  
 aatgtcacca accgacatgt cagtgccgaa aggttaccgt acttaccagt ggaatgtgaaa 1560  
 accattgata tgggcggtac aggttaccgt acttaccagt ggaatgtgaaa 1620  
 gctgtgcgaa atctgcatgt ttatgtcgaa gatgtgcac ttccagccaa tgcatacaaa 1680  
 gatgtaacct actttgaaac gtttaacaaa ggaactacta ctgcatccag ttccaccgcc 1740  
 tatgaaacag ctgatgggtaa cgaagccatc agctccatcg acaaattccg ctctccctac 1800  
 gatccccctg aagctctgga ggaattccgc gtcaccaaac tgaacaacac caactccaac 1860  
 ggtgcggtgc aaaaacgtct attcaggac gccgactatg cgaccgaagt gtcccaacatg 1920  
 ctgttcgaag cgcagtccg gatccagac aactccgtgc tggcaaaagc taaccaggtg 1980  
 tcgaaaagcg agatattcca gcaggccggt taa 2013  
 ccgcagcagg ttctgtctct gctgcagggt

<210> 12  
 <211> 1263  
 <212> DNA  
 <213> Escherichia coli

<400> 12  
 atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaataa tatcaacaag 60  
 aaccagctctg cgtctgcgag ttctatcgag cgtctgtctt ctggcttgccg tatcaacagc 120  
 gcgaaggatg acgcgcgggg tcaggcgatt gctaaccgtt ttacttctaa cattaaaggc 180  
 ctgactcagg ctgcacgtaa ccccaacgac ggtatttctg ttgcgacagc caccgaaggc 240  
 gcgctgtccg aaattaaacaa caacttacag cgtgtctgtg aactgacggt tcaggctctc 300  
 accgggagcta actccgattc ggaactggac cagttcaacg gcgtgaaact gctgtccaaa 420  
 gacgaatttg accgcgtatc cggccagacc aacgagtgcc aaacgattac tatgtatgtc 480  
 gatggctcga tgaataatca ggttggggca aatgattgac agactatcac tattgatctg 540  
 aaggggaaaa ttcaacttca actaaacctc tcaagttttg actgctacaa ctatgggcacc 600  
 aaaaagatg atctcatctac actaaacctc caagtggcag taggtgtctg cgaactttaa 660  
 agtgtttaaag atggggccac catcaataag aaattagtgt agaaagacgg taagtactac 720  
 gataaaactg caggatcggt aggtacccta gatgccgaag tagatactag taagggttaa 780  
 gtaaatgaca ctaaaagtag taagtactac actactccta ctgcagcagc ggaagttaact 840  
 attaaactca acttaccaaa tgaagtgga actactccta ctgcagcagc ggaagttaact 900  
 actgttggcc gogagttaaa attggaatgct tctgacttaa aagcacaaca atcgtctgtc 960  
 gtgtataaag ataaaagcgg caatgatgct tatatactc agaccaaaag tatcaaacat 1020  
 aatcaatcaa ctttcaatgc cgtaataatc agtgatgctg ggtttttatc tattgggtga 1080

tctacaacgc	cgccaagcaa	tttaacagct	aaccgcgtta	agcgtcttga	tgatgcaatt	1020
gcactcgttg	ataaatcccg	ctctctcttc	ggctgcgctc	agaacgcgtc	ggattctcgc	1080
attgccaacg	tgaacaacac	cactaccaac	ctgctctgaag	cgagctcccg	tattcaggacg	1140
gctgactatg	cgaccgaagt	gtccaacatg	tcgaagcgcg	agattatcca	gcaggccggt	1200
aactccgtgc	tggcgaagc	caaccaggta	ccgcagcagg	tctctgtctc	gctgcagggt	1260
ttaa						1263

<210> 13  
 <211> 1368  
 <212> DNA  
 <213> Escherichia coli

<400> 13	agctctctct	gagctccgcc	attgaacgct	tctctcttgc	cctgcgtatt	60
aacaaatctc	aagatgacgc	agcaggctcg	gcgatctgta	acggttttac	agcaaatatt	120
aaaggctcga	ctcaggtctc	ccgtaacgcg	aaigatggta	ttctctgtgc	gcagaccatt	180
gaaggtgcgc	tgaatgaaat	taacacaacac	ctgcagcgtg	taagtgaact	gactgttcag	240
gcaactaacg	gtactaaact	tgacagcgat	cttctctcta	tcacagcgta	aattactcaa	300
cgctctggaag	aaattgacgc	tgatctctgag	caaacctcagt	ttaacgcggt	gaagtcctt	360
gctgaaaaata	atgaaatgaa	aattcagggt	ggcgctaatg	atggtgaaac	catcactatc	420
aaatctggcaa	aaattgatgc	gaaaactctc	ggcctggagc	gttttaatat	cgatgcgcgc	480
cagaaaaagcaa	ctggcagctga	ccctgatttct	aaatttaaa	cgacaggtac	tgataactat	540
gatgttggtg	gtgatgctta	tactgttaac	gtagatagcg	gagctgttaa	agatactaca	600
gggaatgata	tttttgttag	tgacagcagat	gggtcacgtg	caactaaact	tgacacaaac	660
atagcttgta	cagggatgta	tgctacagca	ctcgacagcg	cggtcaagaa	taaacgacag	720
aatgataaat	tcacgtttta	tggagtggaa	ttcacacaaa	caactgcagc	ggatggcaat	780
gggaatgggt	tatatctctg	agaaatgtat	ggtaagtcag	tgacatttac	tgtgacagat	840
gctgacaaaa	aagctctctt	gattacagat	gagcagcttt	acaaaaaatg	cgctggcctt	900
tatacgacaa	ccaaagtgtg	taacaaggct	gccaactctt	ccgatcttga	tctcaatgca	960
gctgaaaaaa	caggaagcac	tgctctctgt	aacaggtgcaa	cttacgatgt	tagtgagat	1020
ggtaaaaagc	taaccggagac	ggtagttggt	aacataaaag	tcagtatctc	gagcaaatca	1080
gaaggtggta	gcccgatctt	ggtaaacgaa	gatgcagcaa	aactcgttga	atctaccacc	1140
aaccgcgtcg	aaactatcga	caaaagctat	gctaaagtgt	acaaactcgc	tctcagcttc	1200
gggtgcagtac	aaaaacggtt	cgactctgct	atcaccaact	tggcaacac	cgtaaacac	1260
ctgtctctct	cccgtagcgc	tatcgaagat	gctgatacgc	cgaccgaagt	gtctaacatg	1320
tctctgcgcg	agatcctgca	acaagcgggt	acctctgttc	tggcgcag		1368

<210> 14  
 <211> 1788  
 <212> DNA  
 <213> Escherichia coli

<400> 14	tcattaatac	caacagcctc	tgcgtgatca	ctcaaaataa	taaccaaacg	60
atggcacaag	cgctgtcgag	ttctatcgag	cgctctgtct	ctggctctgc	tattaacacg	120
aaccagctcg	acgcagcggg	tcaggcgatt	ctcaacctta	gcttaaaaggc	180	
cgcaaggatg	cgcccgctaa	cgccaacgac	ggatctccgc	ttgcgcagac	cacgcgaagg	240
gcgcgtctcg	aaactcaacaa	caactctacg	gctatccgtg	aactgcagct	tcaggctctt	300
acggcgacta	actccgattc	ggatctggac	tccattcagg	acgaataatca	atcccgctctg	360
gacgaattgt	acgcgctatc	tgcccgacgc	cagttcaacc	gogtgaacgt	actggcgaaa	420
gacgggttca	tgaaaaatca	ggttggtgcg	aatgacggcc	agactatcac	gattgtcttg	480
aaagaattgt	actcagatcac	gctggggctg	aattggttta	actgtgaatg	tcccggtacg	540
atagccataa	aagcggcgac	cattagcgac	ctgacagcag	cgaaaaatgga	tgctgcaact	600
aaactcataa	ctacacacaa	taatgcgctg	actgcataca	agggcgttga	tcaactgaaa	660
gatgggtgac	ctgttactat	caaagcagat	gctgctcaaa	ctgccaccgt	ttatacatac	720
aatgctacag	ctgttaactt	ctcaatcagt	aatgtatcga	ataatacttc	agcaaaaagca	780
ggcgatgtag	cagctacgct	tctccgcgcg	gctgggcaaa	ctgctagtgtg	tggtttataaa	840
gcagcaagcg	gtgaagttaa	ctttgatgtt	gatgcgaatg	gtaaaaatcac	aatccggagga	900
cagaaagcat	atttaactag	tgatggttaac	ttactacaaa	acgatgcctg	tggtgcgact	960
ggcgtctacg	ttagtggttt	attcaagaaa	gctgggtgatg	gtcaatcaat	cggttttaag	1020
catctcactc	cagtcacgat	ggggggaaca	acttataact	ttaaaaacggg	gtcgtgatgct	1080
gatgtctgaa	ctgtcaacgc	aggggttatcg	ttcactgatg	cagctagcaa	agaacacggt	1140
gtcaataaag	tgcctacagc	taaacaaaggc	aaagcagttg	cagctgcagcg	tgataactcc	1200
tcaaaatttc	ctctaataac	tggcgttccag	acgtatacag	ctgtatttgc	cgagggtgac	1260
ggtaactgcta	gcgcaaaata	tggcgaataaa	gctgacgttt	ctaatgcaac	agcaacatac	1320

actgatgctg	atggtgaaat	gactacaatt	ggttcataca	ccacgaagta	ttcaatcgat	1380
atggcaaacg	cgaggtgaac	tgttgattct	ggaactggta	cggttaataa	tgccgcgaaa	1440
gtaggggctg	aagtatatgt	tagtgctaatt	ggtaacttaa	caacagatgc	aactagcgaa	1500
ggcaccagtaa	caaaagatcc	actgaaagct	ctggaatgaag	ctatcagctc	ctatcgacaa	1560
ttccgtttctt	ccccgggtgc	tatccagaac	cgctggtgatt	ccgcagtcac	caacctgaac	1620
aacaccacta	ccaaactgttc	cgaaagcgag	tcccgtrattc	aggacgcgca	ctatgcgcagc	1680
gaagtgccca	acactgtcgaa	agcgacagtc	attcagcagg	ccggttaactc	cggtgctggca	1740
aaagccaacc	aggtaccgca	cgaggttctg	tctctgctgc	agggttaa		1788

<210> 15  
 <211> 1653  
 <212> DNA  
 <213> Escherichia coli

atggcacaag	tcattaatac	caacagcctc	tcgctgatca	ctcaaaataa	tatcaacaag	60
aaccagttctg	cgctgtcgag	ttctatcgag	cgctgtgctt	ctggcttgcg	tattaacagc	120
gcgaaggatg	acgccgcagg	tcaggcgatt	gcctaacggt	ttacttctaa	cattaagaagc	180
ctgactcagg	ctgcacgtaa	cgccaacgac	ggatattctg	tgccgcagac	cactgaagct	240
gcgctgtcgc	aaatcaacaa	caacttacag	cgctgtcggt	agctgacggt	tcagggtctc	300
accgggacta	actccgatcc	tgacctggac	tccatccagg	acgaaatcaa	gctctgctgt	360
gacgaatttc	acgcggtatc	cggtcagacc	cagttcaacg	gcgtgaacgt	gctggcgaaa	420
gacggttctg	tgaaaaattca	ggttgggtgc	aatgcaggcc	agactatcac	gattcgactg	480
aagaaaaattg	actcagatgc	gctggggctg	agtgggttta	atgtgaattg	tgccggggct	540
gttgtcaaca	ctgctgcatt	taaagctgac	ttggtagctg	ctaattgcaac	tggtgtaggc	600
aacaaatata	ctgtgagtcg	gggttcagac	gctgctaaag	cgctgtgatt	gctggctgga	660
gttagtgatg	gtgatactgt	tcaggcaacc	attaaataacg	ctgtcggaaac	gcggcgtaagt	720
gcacagcaatt	acaagtatga	cagtgcaagt	aagttctact	cttttgatac	caacaacgct	780
tcagctgcgc	atgttcagaa	atatgtgac	ccggcggttg	gtgataccgc	taagggrcat	840
attactatcg	atggttctgc	acaggatggt	cagatgacga	gtgatgttaa	aatattcgca	900
agcaatggag	ataaaactta	cattgatata	actggcgctc	taacgaaaaa	cggttttagt	960
gcttcttttg	ctgaggctag	ctctctcttt	cttgctcaaca	cttgacgcca	agcgacaacc	1020
attgacatcg	gcggtacctc	tatctccttt	caaggcgctt	tcgataaagc	tgtatcaaac	1140
actttatcag	taacaggctg	aaaagttgat	gggttatagc	tgacacggcc	aactggcgct	1200
tcctggaaaag	atgtgtgatt	cactacgcga	ggttatagcg	gataaacaac	ggcggttgac	1260
gtaacaaaaag	gttgtgtctc	gggttatatt	ggttatatac	gcagtgatgt	caatctctat	1320
actgtcagatt	tttatctaca	tgatgatggt	tcagtgacta	acggcagcgg	taaggcgact	1380
tataaaagat	ctgacggtaa	attgacgaca	gatgctgaaa	ctaaagctgc	aaccaccgcc	1440
gatccccctga	aagctctgga	cgaaagccatc	agctccatcg	acaaattccg	ctctccctcc	1500
ggctgcgtgct	agaacgctgc	ggattccgcg	gtccaccaac	tgaacaacaac	cactaccaaac	1560
ctgtctgaag	cgagctccgc	tattcaggac	gctgactatg	gcaccgaagt	attccaacgt	1620
tcgaaagcgc	actcatccca	gcaggccggt	aactccgtgc	tgccaaaagc	taaccaggta	1680
ccacagcagg	ttctgtctct	gctgcagggt	taa			1740

<210> 16  
 <211> 1689  
 <212> DNA  
 <213> Escherichia coli

atggcacaag	tcattaatac	caacagcctc	tcgctgatca	ctcaaaataa	tatcaacaag	60
aacacagctg	cgctgtcgag	ttctatcgag	cgctgtgctt	ctggcttgcg	tattaacagc	120
gcgaaggatg	acgccgcagg	tcaggcgatt	gcctaacggt	ttacttctaa	cattaagaagc	180
ctgactcagg	ctgcacgtaa	cgccaacgac	ggatattctg	tgccgcagac	cactgaagct	240
gcgctgtcgc	aaatcaacaa	caacttacag	cgctgtcggt	aactgacggt	tcagggaacc	300
accggtacca	actcccgatc	tgacctggac	tctatccagg	acgaaatcaa	atcccgctgt	360
gacgaatttc	acgcggtatc	cggtcagacc	cagttcaacg	gcgtgaacgt	gctggcgaaa	420
aagagttctca	tgaaaaattca	ggttgggtgc	aacgatggcc	agacccatcac	tatcgactgt	480
gaagaagatg	actcttctac	cttgaacctg	acaggtttta	acgtttaacg	ttctggtctt	540
ggcggaata	ctgcagcaac	taaagctgat	ttaacggctg	ctcaactctc	tgaccggggt	600
gcgcagacag	caaatggatc	agttacttat	actgtcagtg	ctggttataa	agaatccact	660
ctgcagcagc	ttatgtctag	catcaaaagc	ggcagtgctc	cgacttctgc	attactactga	720
accattaata	atggctctgg	tgattccagt	gcgctgactc	ccaatgacta	tacttatgac	780
ccagcaaaag	gcgactccac	ttacgacgta	gcttcaacgc	ccaataatac	tgctgccagc	840
gttcagttct	tcttcagccc	gaaagcagg	gataccgcaa	atctgaaagt	aacggttggt	900

acgacatcgg ttgatgtcgt tctggccagt gatggttaaga ttacagcaaa agatggttct 960  
gcattatata tcgacagtac aggttaacctg actcagaaca ctctcaacct gacctctgct 1020  
aaactggccta cctcgactgg ccttcagggg tctgggtgtg ctccaacctc cactactgaa 1080  
gatggcacta atattgatat tgctgctaac ggttaatttg ttactgctgg tactggctatc 1140  
agtgtgactt ctctgcagtc agcgactaaa tctacgggct ttaactgtgg tactggctatc 1200  
acaggtctcga cagtaggtac tgaagtataa gtgactatcg geggactac tctgcagtc 1260  
tacaccagca aagatgggttc cctgactact gataacacca cttaaactgt tctgcagaaa 1320  
gatggctctg taaccaacgg ttcaggtaaa ggggtctatg cgtgaaagc cctggtatgag 1380  
actaccagcg ctgcaaccaa agcgcaaccc accacggatc ctatcagaa ccgtctggat 1440  
gcaatcagcc agatcgataa gtctcgttca tccctgggtg ctgaagcgca gtccgttat 1500  
tcgcggtctca ccaacctgaa caacaccact accaactctg ctgaagcgca gtccgttat 1560  
caggacgccc actatcgac cgaagtgttc aacatctcga aagcgagat gctctgctg 1620  
gcggttaact cgtgctggc aaaagccaac caggtaccgc aacaggttct gctctgctg 1680  
cagggtctaa 1689

<210> 17  
<211> 915  
<212> DNA  
<213> Escherichia coli

<400> 17  
gcgctgtcga cttctatcga gcgctctctc tctggtctgc gtattaacag cgctaaagat 60  
gacgtctcgg gccaggcgat tgctaaccgc ttcaattcta acatcaaaag tctgactcag 120  
gcgcgacgta acgccaacga cggattttct ctggcgcgaga cggctgaaag cgcgctgtca 180  
gagatttaaca caaacttgca cggatttctg gaactgacgc ttccaggcgc taccggcagc 240  
aaactctgatt acgacctgtc ttctattcag gacgaaatca aaacccgtct tgatgaaatt 300  
gacgcgtgat cgggtcagac ccagttcaac ggtgtgaagc tctgtcgaa aaacgactcg 360  
atgaagattc agatttggtc caatgataac cagacgatca gcatgtgctt gaaacaaatc 420  
gacagttacca ctttgaatct gaaggattt accgtgtcgc cagacgatca ttcacgctcg 480  
gcgaaactga cgggtcgtga ttgtacagca atgtgtctg cggatgtcaa ggaatgtcgg 540  
ggtaaaacaa tcaatttact gtcttcaact gacacccgct tagtcaattc cagtagcgcg 600  
gtcgttgatt ctgcaacggg taaatacatg gcagccactg cgcggagcgc accgttaaaa 720  
gcggcggtga cttgttggtc aaacggagtg gcgggagcgc cctccctcgg tgcgctcaa 780  
gcactggagt ccgaactcgc taaagtgcac aaattccgct caaccaacct gctcgaagcg 840  
aacgcgtctg attctcggtt caccacacct aacaacacca ccaactatgc gaaagcgag 900  
cagttccgta ttacaggacg cagctatcgc accgaagtgt 915  
attatccagc agggc

<210> 18  
<211> 1665  
<212> DNA  
<213> Escherichia coli

<400> 18  
atggcacaag tcattataac caacagctc tgcgtgatca ctcaaaataa tatcaacaag 60  
aacagactcg cgtgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120  
gcgaaggatg acgcgcaggg tcaggcgatt gctaaccgtt ttactcttaa tattaaaggc 180  
ctgactcctg ctgcacgtaa gcaccaatga ggtatttctg ttgcacagac cactgaaggc 240  
gcgctctcgg aaatcaacaa caacttacag cgtattctgt aactgacggt tcaggccact 300  
acaggggacta acrcgacttc tgacctggac tccatccagg cagttcaacg gctgtccaaa 420  
gacgaaattg acccgctgat cgggtcagacc cagttcaacg gcgtgaagct gctgtccaaa 480  
gatgggtcaa tgaaaattca ggtcgcgca aatgatgttg aaacctacac gattgatctg 540  
acagccaata ctgctgcaac acttaaaagt atggttggtt taaaactcga taatcggggg 600  
gtccactacg cgtgagttaa tagataatt gctgacaaag ccgtcgcaag tagcagcgat 660  
tttttgatag cgtgagctgg tgttgatggc agtaaaagt ttccagggagc tagctgttgt 720  
tattgtcgag cgtccctcgg tacgcacggt gaataactt atcataaaga tactaacaac 780  
tatagcgctt cgtcttcagt tgatcgact caactggcgt cattctgtaa tctcgaagcg 840  
gatgctacca ctgctgcaac agtaagatt ggcaacgcta caacagctca agagcaaaaa 900  
gtcatctatg ctaaaagatg ttctttaaact cgtctgtagt accgctgcgc tctctattct 960  
gatgatactg taaacttaag taaaactaac gcaggcactg atactcaagc taaactgtct 1020  
gacttaatgg caaacatgc taatgccaaa acagctatta caacagataa aggtacattt 1080  
actcgtataa cgacaaagtt tgatggggtg gatatttctg ttgatcttc cagctttgat 1140  
aacgcgctga aaaaagagac atgatttggt ttaccttacc tgcagacagc gattcagac 1200  
acagctaaata atggcactgc tgcactcagc tatttagctg atggaaaagt gagcaaaact 1260

cctgcgcgagt	atttttgccta	agctgatggc	actattacta	gtggtgaaaa	tgccgctacc	1320
agtaaaagcta	tctatgtaag	tgccaatgg	aacttaacga	ctaatacaac	tagtgaatct	1380
gaagctacta	ccaaaccgct	ggcagcattg	gatgaacgta	tcgcgtctat	cgacaatttc	1440
cgttctctccc	tggtgtgctat	ccagaacgct	ctggattccg	cagtaaccaa	cctgaacac	1500
accctaccaca	agctgtctga	agccagctcc	cgtattccag	acgcgcagta	tgccagcgaa	1560
gtgtccaaac	tgctgaaagc	gcagatcatt	cagcaggccg	gtaactccgt	gctggcaaaa	1620
gccaaaccagg	tacgcgcaga	ggtctctgtc	ctgctgcagg	gttaa		1665

<210> 19  
 <211> 1842  
 <212> DNA  
 <213> Escherichia coli

<400> 19						
atggcacaag	tcattaatac	caacagcctc	tcgctgatca	ctcaaaaata	tatcaacaag	60
aaccagctcg	cgcgtgcgag	ttctatcgag	cgtctgtctt	ctgcgttgcc	tattaaacgc	120
gcgaaggatg	acgcgcgagg	tcaggcgatt	gctaaccggt	ttactttctaa	cattaaaggc	180
ctgactcagg	ctgacagtaa	cgccaaacgac	ggattttctg	ttgcgcagac	cactgaaggc	240
gcgctgtccg	aaattaaaca	caacttacag	cgtattccgt	aactgacggt	tcaggcgacc	300
accggaaacta	actccacctc	tgacctggac	tcactccagg	acgaaatcaa	atcccgctct	360
gacgaaattg	accgcgtatc	tggtcagacc	caggttcaacg	gcgtgaacgt	gtctgtctaaa	420
gtaggctcga	tgaaaattca	ggctcggcgcg	aacgatggcg	aaacgattac	tattgatctc	480
aagaaaaattg	actctgatac	gctgaattctg	gctggtttta	acgttaacggt	taaaaggctct	540
gtagcgaata	cgcctgcgac	tacagataat	ctgaacttgg	ctggttttac	agcgggatct	600
aaagctgtgtg	atggcacctg	aactttatag	aaaaattgct	agtttgcgcg	cagcagctgca	660
agcaactgtac	tggtctgctgc	taaaagtggc	gacgaattta	cgttcgctgg	taataacggc	720
acagggtatag	ctgcgaactgg	ggggacttat	acttatcata	aggactctaa	ctcatacagg	780
tttagcgcaa	cggctgcctc	taaagattct	ctgttgagca	cactggcacc	aaacgctggc	840
gatacatttta	ccctcaaatg	gactatttgt	tcataaactc	aagaagttaa	cgttagcaaa	900
gatggtcaga	tgactatcag	cgatgtaag	gcgctgtatt	tagatgagaa	gggcaacctg	960
accocaaacag	gtagtggcac	aaccaaggct	gcacacctgg	ataactgat	ggccaataca	1020
gatactacag	cgaagaatgc	ctatgttaac	ctgcggcgac	cagctgttgg	gacagtaatc	1080
gaagcaaaaag	gaatgacct	cactctgcgt	gggtgtaag	ctcagtgct	aaaagacgct	1140
gccttaaatg	cgcgatatgc	gacctcaaat	actactggta	ctccgggtga	tgccgggacc	1200
gcggggagcgc	ctgcgaactgc	gggtaatgac	gcgctggggc	acgcgctgct	agcggagctc	1260
gataataacga	cggcagatgt	tgccgatata	tcattctcag	cttcgcaaat	ggcgagcatt	1320
cttcagagata	aagatttca	cttaagtgt	gttagtgata	cttacaacgt	gaccgcaaat	1380
gcgtgtacta	tcaatggcaa	agcagcaaac	attgatgaca	gcggcgcaat	caacagccaa	1440
accagtaaaag	ttgtcaatta	tttgcctcat	actaacggt	cgtgactaca	cgatccaggc	1500
tcactatttt	atgcgcagaca	agatgttagc	ctgaccacgc	atgcagcaac	caaagccgaa	1560
accacgcgcgc	atccctcgaa	agctctggac	gaagccatca	gctccatoga	caaatccgcg	1620
tcctccctcgc	gtgcggtgca	aaacgctctg	gattccgcgc	tcaccaacct	gaaccaacc	1680
accocaaacc	tgcttgaaag	gcagctccgt	attcaggagc	cgcactatgc	gcgcaaggtg	1740
tcacaaactgt	cgaagcgca	gattatccag	caggccggta	actccgtgct	ggcaaaagct	1800
aaccaggtag	cacagcaggt	tctgtctctg	ctgcagggtt	aa		1842

<210> 20  
 <211> 1731  
 <212> DNA  
 <213> Escherichia coli

<400> 20						
atggcacaag	tcattaatac	caacagcctc	tcgctgatca	ctcaaaaata	tatcaacaag	60
aaccagctcg	cgcgtgcgag	ttctatcgag	cgtctgtctt	ctgcgttgcc	tattaaacgc	120
gcgaaggatg	acgcgcgagg	tcaggcgatt	gctaaccggt	ttactttctaa	cattaaaggc	180
ctgactcagg	cgcgcgcgta	cgccaaacgac	ggattttctg	ttgcgcagac	cactgaaggc	240
gcgctgtccg	aaattaaaca	caacttacag	cgtgtgcgtg	agctgactgt	tcaggcgacc	300
accggtaccga	actccagctc	tgactcggac	tcattccagg	acgaaatcaa	atcccgctct	360
gacgaaattg	accgcgtatc	cgttcagacc	caggttcaacg	gcgtgaacgt	gtctgtctaaa	420
gcagcgttcca	tgaaaaattca	gggtcggcgcg	aatgatggcc	agaccatcac	tatcgactgc	480
aagaagaattg	actcttctac	gttgaaactg	actggtttta	acgtgaaattg	ttctggtctc	540
gtggcgcaata	cgcgcggcac	taaaagcggat	ttggctgctg	ctgcaattgg	tacccctaggc	600
gcagcagatt	ctacaggtgc	cattgcttat	acagtaagtg	ctgggctgac	taaaactaca	660
gcgcgcagatg	tgtgtctgag	cctcgctgat	ggtagcacta	ttacagccac	agcggcgaaa	720
aatggtcttg	ctgcaggagc	cacttccaat	gcctataaac	ttacaacaga	taataataca	780

```

tttacttatg acacgactgc tacgacagct gagctgcagt cttacctgac tccgaaagcg 840
ggcgacactg caacatttcag tgttgaattt ggtgttacta cacaagacgt cgtgctgtcc 900
ttagcttgga aactcactgc taaggatggc tctaagcttt acatttgatac aactggtaat 960
ttagctctga acaaaaatgg taacgggtgtt ggaacactcg cggaaacgac tctgagtggt 1020
acttcgattg tactgaatgg ttcaacggct gttaaatcca caattactac agctgataac 1080
atgtctgttt caggcgctgt aattagttca gctgctctgc aatctgcaag caaaaagact 1200
ggtttcactg ttggtacagt agacacagct ggttatatct ctgtaggtac tgaagggtacg 1140
gttcaggcat atgatctgct gacttctggc aaaaaagctt cttacaccaa cactgacggt 1320
acactgacta ctgataaacac cactaaactg tatctgcaga aagatggctc tgttaaccaac 1380
ggttcaggta aagcgggtcta tgtagaagcg gatggtgatt tcaactccga cgcgtcaacc 1440
aaagcgcgcga ccaccaacgga tccgctggcc gctctggatg agcgaatcag ccagatcgac 1500
aagttccgtt catctctggg tgcataccag aaccgtctcg attctgcagt caccacactg 1560
aacaacaccca ccaccaacct gtcgtgaagc cagtcoccta ttcaggagcg cgaattatcg 1620
accgaaagtg ccaatatgtc gaaagcgcag atcatccagc aggcgcgtga ctcgctgctg 1680
gcataaagcca accaggtacc gcagcagggt ctgtctctgc tgcagggtta a 1731

```

```

<210> 21
<211> 1380
<212> DNA
<213> Escherichia coli

```

```

<400> 21
aacaaatctc agtcttctct gagctccgcc attgaacgtc tctctctctg cctgcgtatt 60
aacagtgtga aagatgacgc agcaggttcag gcgattgcta accgttttac agcaaatatt 120
aaaggtctga ctcaggcttc cgttaacgcg aatgatggta tttctgttgc gcagaccact 180
gaagggtgcg tgaatgaaat taacaacaaac ctgcagcgta ttctgtaact tctgttccag 240
gcaactaacg gactcaactc tgacagcgat cttctctcta tccaggctga aattactcaa 300
cgctcgggaag aaattgacgc tglatctgag caaacctcagt ttaacggcgt ccaatgctct 360
cgtgaaaata atgaaatgaa aattcaggct ggtgctaatt atggtgaaac catcactatc 420
aattcggcaa aaattgatgc gaaaacatct ggcctggagc gttttaatc cgaatggcgcg 480
cagaagaaga cggcgagtga cctgattctc aaatttaag cgacaggtac tgaataatt 540
caaatatacg gtaactgata ctatactggt ttggtgata tggcgtagt acaggataaa 600
gatggcaaac aagtttatgt gactactgcg gatggttcac ttacgaccag cagtataact 660
caattcaaga tcatgcaaac taagcttgca gtggtctgta aagatttagc tcaagggaat 720
aagattgtct cgaagggtat cgaatttaca atgaccggca ctgtcgctat agatgccaaa 780
ggtaatggta aattaacgcg caatgttgat ggttaaggct tgaaltcac tatttcgggg 840
agtactgata catcagggtac tagtgcaacc gttgcoccta cgacagccct atacaaaat 900
agtgacgggc aattgactgc aacaaaagtt gaaaataaag cagcgacact atctgactct 960
gatctgaacg ctgccaaaga aacagggaag acgttagtgt ttaacggtgc aactacgat 1020
gttagtgcaag atggttaaac gataacggag actgcttctg gtaacaataa agtcactgat 1080
ctgagcaaat cagaagggtg tagcccgatt cgtgtaaacg aagatgcagc aaaaactgtg 1140
caacttaccg ccaacccgct cgaactatc gcacaaagct tggctaaagt tgacaactcg 1200
cgctctgacc tegggtcgct acaaaacgct ttcgactagt ccatcaccaa ccttggcaac 1260
accgtaaaac acctgtctc tgcocgtagc atgctgacta cgcagcagaa cgcgaaggt 1320
ggtcttaaca tgtctctgct gcagatcctg caacaagcgg gtacctctgt tctggcacag 1380

```

```

<210> 22
<211> 1767
<212> DNA
<213> Escherichia coli

```

```

<400> 22
atggcacaag ctattaatac caacagcctc tcgctgatca ctcaaaaata tatcaacaag 60
aacacactcg cgtctgcgag ttctatcgag cgtctgtctt ctggtctgct tattaacagc 120
gcgaaggatg acgcagcggg tcaggcgatt gctaaccggt ttactcttaa cattaaaggc 180
ctgactcagg cggcagctaa gcaccaagac ggtatctctc ttctctctc caccgaaggt 240
gcgtgctctg aaatcaacaa caacttacag cgtgtactgt aactgcacct tcaggcaacc 300

```

accggtacta actccgactc cgcctgggct tctatccagg acgaaatcaa atcccgctcg 360  
 gatgaaatcg acccgctatc tggctcagact cagttcaacg cgtgaaactg cgtggcaaaa 420  
 gagcgtttcca gacgaaattca ggtagggtgct aacgacgcgc agactatcac tatcgacctg 480  
 aaaaaaatcg actctcgatc tctggggctg aatggtttta acgtgaaatg tctcgggacg 540  
 attaccaaca aagcagcaac tgtcagtgat ggtactcgcg caggcggtac attggtgaat 600  
 ggtgcctatg atataaaaac cacttaacac ggcctgacta caactgatgc ctcgcgaaaa 660  
 ttgaatgatg gtgatgttgt tactatcaat aatggtaagg atactgccta taataataat 720  
 ctgcctacag gtgggtttac gacgggatgc tccatctccg gggatcctac cgtcgtgcac 780  
 gctactgccta ataaaaactgc cgtgatgaca ctgcggcgct ctttaacatg tgagcggggg 840  
 aaaaactgtta atgggtctctg gactcagaaat gatggtaacg gcttatgtag taaaaattga taccgatgcg 900  
 gatggtaaga tttctattgg tgggtgtgct gactatcgac acaactacgc atttggttac tctgcgtgca 1020  
 actaaccagcg caggatgtac atctcaagga gactacacgc atttggttac tctgcgtgca 1080  
 tctgcctatg gtaagggtgg atccctgacc ttgggtgaca cgcgtataaa aattggctag 1140  
 ggtcagcgtg ggggtgatcc tgcagcgcct tgcagtatg tactgggacg cgtcgtgcga 1200  
 tctaaatcag taagcaaggga tctgttctct gctgatacta aagcaactcg taacacgaca 1260  
 acagttgatt tcaactccgg tatcatgact tcaaaaggta gtttcgatgc aggtacatca 1320  
 actgatcatc tcaaaagatgc agatgggtgct atcaccaaaa ctaaagaata ccgactctg 1380  
 tatgctgttaa ataaagatgc tgggtgaagt accggttctg attatgctgc gactgatgcg 1440  
 ccgcataaag ctgtgtgatga tactaaatat aaacgcgacta tccggcgacg agttaaacctg 1500  
 aattctcgacg gtaaatgtac cactgatacc atcgacaaat tccgcttcac cctgggtgct 1560  
 ctggcgtgcc tggagcgtgc tatcagctcc cgcagtcacc atcgcaaaa acacactac 1620  
 atccagaacc cctgtgatcc cgcagtcacc atcgcaaaa acacactac 1680  
 gaagcgcact ccgctattca ggcgcgcgac tatgcgacg aagtgctcaa catgtcgaaa 1740  
 ggcgcagata tccagcggc cggtaactcc gtgctggcaa aagccaacca ggtaccgcag 1760  
 cagggtctctg ctctgcata gggttaa 1767

<210> 23

<211> 1383

<212> DNA

<213> Escherichia coli

<400> 23

acaaaaaaac agtctcgctc tgcgaactct atcgagcgcc tttctctctg tctgcgttat 60  
 aacagcgctga atagatgacg tgcggggccag gcgattgcta accgcttcac tcttaaacatc 120  
 aaaggtcttga ctccggcgcg agtaaaacgc aacgacgcta tttctctctg cgcagccact 180  
 gaaggcgccg tgcctgagat taacaacaac ttgcagcgtg tgcgtgaatt gactgtacag 240  
 gcgacgacgc ggaactaaact tgcattctgac ctgtctctca tccaggatga aatcaaatcc 300  
 cgttttaagcg aaattgacgg tgcattctggt cagactcagt taacggcgct gaacgtaactg 360  
 gctaaagatg acaacctgtc tattcaggta ggtgcgaatg acggtcagac tatcaaatat 420  
 gactcgcagc aaactgattc tcatcacact ggtctggatg gttcagcgt taaaaataat 480  
 gatcgcagta aaacagatgc tgcgctgaat actcttgggg gggggggcag tctgtgtgct 540  
 gtgcgactcg caacaacacg tttgactgct atcactggtc tccgtagcgg tgcattcagc 600  
 gaaattgcta aagacgataa tgggtgattac tacgcgcatg tccaggggac tacgggtaat 660  
 actgctgatg gttactatgc tgcgcgatac gacaggtacg ccaacgcaga caagcacata tgacttcaca 720  
 gatggttaac tagatacacc gacaggtacg gatgctgaca cagcgcggtat gactcctggt 780  
 gacgctggct aaacgctttc ctttggcact gatgctgaca cagcgcggtat gactcctggt 840  
 gctctctctg ttaaacttca ggaatgagaa ggcaatgata ctgactacta tgcaatcaaa 900  
 gcacaagatg gcagcctgta tgcgcgcaac gttgatgagg ctacgggtaa agtcaactgt 960  
 aaaaacgcga gctatactga tgcgcgacgc aaagcagta cgcgtgccgc tgtaaaaactg 1020  
 ggtggtgaca atggcacaac cgaattgttt tgcgatgctg cgtcaggtaa aacttacgat 1080  
 gctggtgcaac tgcaaaacgt tgcattctcc agtgcaacca acacggtaac cgcatacccg 1140  
 aacggtataaa caacgcttcc gctggctgccc cttagcagca caatcacgca gatcgcaaaa 1200  
 ttcgctcact cccctcggtgc ggtgcagaa cgtctggatt ccgcggtcac caacctgaac 1260  
 aacacaccta ccaacctgtc tgaagcgcag tcccgatct aggaacgctga ctatggacac 1320  
 gaagtatcca acatgtcgaa agcgcagatc atccagcagg caggtaactc cgtcgtgtcc 1380  
 aaa 1383

<210> 24

<211> 1197

<212> DNA

<213> Escherichia coli

<400> 24

cgcgtctgca cttctatcga ggcctctctc tctggtctgc gcattaaacg cgtcaaatg 60  
 gacgctcgcc gccaaagcat tgcataacgc ttcacttcta acatcaaaag tctgactact 120  
 gcgcgacgta gcagccaaga cgttatctct cgtggcgaga ccaactgaag cgtactgtct 180  
 gaaatcaaca acaacttgcg cgtgtgttgt gaaactgacg ttcaggccac taccggtact 240



aactctgatt	ctgacctgtc	ttcaatacag	gacgaataca	aatccggtct	cgatgaaatt	300
gacgcggtat	ccggtcgagc	tccgttcaac	ggcgtaatag	ttctttccaa	agatggttca	360
atgaaaattc	aggttgggtg	gaatgatgg	caactatct	ccatcgatct	gaagaaaatt	420
gatttcttca	ctttggggct	gaatggcttc	tcagtttcta	aaaactctct	taagtcaag	480
aatgctatca	catctatccc	gcaagccgct	agcaatgaac	ctgttgatgt	taacttcggt	540
gatactgacg	agctcgcagc	aatccgagcc	aaattggggg	tttccgatcc	gtcaagctcg	600
tcgctgcaca	acatccttga	taaagatgg	aaggcaacag	ctgatgatgt	tgctcagctc	660
ggtaaaagct	tctatgtctg	ttctgttaat	gccggttcag	gtaaaagtaac	cttaaacacc	720
attgatgtta	cttatgatga	ttatgcgaac	ggtgttgacg	atgccaagca	aacaggctcag	780
ctgatcaaa	tttcagcaga	taaagcggcg	cgagctcaag	gttttgcac	acttcaaggc	840
aaaaactatt	ctgcgtgtga	tgcggcagac	at tcttaaga	atggagcaac	agctcttaag	900
tttaactgat	tgaatttaag	tgtatgttact	gatactaagt	gtaaggttaac	caacaactcg	960
actgagcaat	ttgaagggtg	ttcaactgag	gatccgctgg	cgcttctgga	taaaagtatt	1020
gcatacgtcg	acaaattccg	gtcttctcta	ggtgcggtgc	agaacccgtc	cgatcttcgt	1080
atcaccaacc	tgacaacac	caccaccaac	ctgtctgaag	cgagctccg	tattcaggac	1140
gcgcactatg	gcagcgaagt	gtccaacatg	tcgaaaagcg	agatcatcca	gcaggca	1197

<210> 25

<211> 1674

<212> DNA

<213> Escherichia coli

<400> 25

atgggcacag	tcattaatac	caacagccctc	tcgctgatca	ctcaaaataa	tatcaacaag	60
aaccagctctg	cgctgtcgag	ttctatcgag	cgctgtgctt	ctgggttgcg	tattaaacag	120
gcgaagagatg	acgcgcgagg	tcaggcgatt	gctaaccggt	ttacttctaa	cattaaaggc	180
ctgactcagg	ctgcacgtaa	gcaccaacac	ggatattctg	ttgcacagac	caactgaaggc	240
gcgcgtctcg	aaatcaacaa	caacttacag	cgatctcggt	acgtgacggt	tcaggccaat	300
acaggggacta	actccgattc	tgacctggac	tcactccagg	acgtatactaa	atctctgtct	360
gacgaaattg	acgcggtatc	tggtcagacc	cagttcaagc	gggtgaacgt	ggtgtctaaa	420
gatggctcga	tgaaaaattc	ggctggcgcg	aacgatggcg	aaagcatatc	catttgatctg	480
aagaaaaattg	actctgatcat	gctaaaactg	gctggtttta	acgtgaatgg	ggctggctct	540
ttgtgacaatg	ccaagcgac	ctataagat	ctctgtgatg	ctgggttttc	tcaggcgcaa	600
gctgatgctta	atggcaaaat	cacttatacc	aaagacacg	ttactaaaat	cgacaagcgc	660
acagcggctg	atgtatttgg	caaaagcgct	gctggcgata	gcattacctc	tcggggcact	720
gatactggct	taggagtcgc	tgtctgtgcc	tgcacttaca	ctcaaatgac	agccaataag	780
ttctaacatt	ttgatgtctac	tggtgttgcc	aaagcgagtg	ctggaacggc	actgaaaggg	840
tacttagggc	catctaacac	cggtaaaaat	aatatcggtg	gtacogagca	agaagttaac	900
ggcaactctaa	atggctccat	caccgatcac	aatggcgatg	cgctgtatct	cgatagtaac	960
ctgttttgtg	gtgctcagga	gcgaatttg	ggsgctgctg	ataaagcaac	tgataataaa	1020
gatcaaaactg	ctggatccgt	tgatttcaaa	ggcgcgctta	tttctgtctg	gtcaatggca	1080
tcaacctctaa	ataatgggtc	ctatacagcc	aacgtaggtg	gtaaaggctc	tgccgttaac	1140
gtggcgcgac	ctcagacagg	tgccgcagat	gtgtataaag	ataccactgg	cgactcagcg	1200
actgaagatg	acgaaacggt	taccgcgacc	tactacgggt	ttgtgtgatg	taaaagtctt	1260
gacggtgaa	gttctactgt	ctataaagct	gatgatgggt	ccatcactaa	agatgcactc	1320
acgaagtctg	aagcaaacac	tgacctctgt	aaagcccttg	acgacgcaat	cagccagatc	1380
gacaaaactct	gctctctcct	cggtgcgctt	caaaacccgt	tggattcccg	cgctaccaca	1440
ctgaacaaca	ccactaccaa	ctgtcttgaa	gcgcagtcoc	gtattcagga	cgccgactac	1500
gcgacccaag	tgtccaaat	gtcgaagcg	cagatcatcc	acaggccgcg	taactccgtg	1560
ctggcaaaa	ccaaccaggt	accgcagcag	gttctgtctc	tgctgcaggg	ttaa	1620

<210> 26

<211> 1365

<212> DNA

<213> Escherichia coli

<400> 26

aacaaatctc	agtcctctct	tagctctgct	attgagcgct	ttctttctcg	cctgcgtatt	60
aacagtgcta	aagatgaacg	agcaggtcag	gcgatgtcta	acggttttac	ggcaaatatt	120
aaaggtctga	ctcaggcttc	ccgtaaacgc	aatgatggta	tttctgttgc	gcagactact	180
gaaggtgcgc	tgaatgaat	taacaacaac	ctgcagcgct	taagtgaact	gactgttcag	240
gcacactaac	gtactaactc	tgacagcgat	ctttcttcta	ttcagcgaga	aattactcaa	300
cgctcggaa	aaattgacg	tgtattctgag	caaaactcagt	ttaacgcggt	gaagctctct	360
gcgcaaaaat	atgaaatgaa	aattcaggtt	ggtgctaagt	atgggggaaac	catcactatc	420
aatctggcaa	aaatttgatg	gaaaactctc	ggcctcgagc	gcttttaaat	cgatggcgcg	480
cagaaagcaa	ctggcagtg	ctgtatttct	aaattttaa	cgacaggtac	tgataattat	540

caaatataacg	gtactgataa	ctatactggt	aagttagata	gtggagcagt	tcaaaatgag	600
gatgggtgacg	caatttttgt	tagcgctacc	gatggtttct	tgactactaa	gagtgataca	660
aaagtccggtg	gtacagggtat	tgatggcgact	gggtcttgcaa	aagccgcgagt	ttcttttagct	720
aaagatgcgct	caatataaata	ccaaggtatt	actttccaca	acaaagcgac	tgatgcattt	780
gatggcgagtg	gtaacggcgac	tctaaccgct	aaatattgatg	gcaaaagatg	aacctttact	840
attgatgcgga	cagggaaggga	cgcacacattt	aaaaacgtctg	atcctgttta	caaaaagtgt	900
gcagggtcagt	tcaactacac	taagggttgaa	acaaaagcgtg	ctacacagc	ggatctggag	960
ttaataaacyg	ctaaaaaagt	gggtagttct	ttagtgtgtaa	atggcgtgaa	ttatgaaagt	1020
agcgctgagtg	gtagacagtg	aactgggctt	ggcaaaaacta	tgtattctgag	caaatcgaaa	1080
gggtggtagcc	cgattctcgtg	aaaagaagat	gcagctgaac	atctgcgttc	tgacctcgtg	1140
ccgctcgaaa	ccatcgacaa	ggcatggct	aaagtgtgaca	gcacacccgt	gaacaaacctg	1200
gcagatcaaaa	acggtttcga	ctctgtctatc	accaaactgt	cgaagaagtc	taacatgtct	1260
tcttctgcgc	gtagcgtgat	cgaagatgct	gactacgcga	cgcag		1320
cgtagcgaga	tcttgcaaca	agcgggtacc	tcgtctctgg			1365

<210> 27

<211> 1740

<212> DNA

<213> Escherichia coli

<400> 27

atgggcacag	tcattaatac	caacagcctc	tcgctgatac	ctcaaaataa	tatcaacag	60
aaccagctctg	cgctgtcgag	ttctatcgag	cgctctgctt	ctggcttgag	tattaaacgc	120
gcgaagagatg	acgcgccgag	tcaggcgatt	gctaaccggt	ttactttcaa	cattaaaggc	180
ctgactcagg	ctgcacgtaa	cgccaacgat	ggatattctg	ttgcacagac	cactgacagc	240
gcgctgtccg	aaatccacaa	caacttacag	cgatcccggt	aactgacgct	caaggctct	300
acccgggacta	atcccgattc	ggatctggac	tcacttcagg	acgaaatcaa	atcccgctct	360
gacgaaatgtg	acgcgctatc	tgcccgagcc	cgattcaacg	gcgtgaaagt	atggccgaaa	420
gacggtttcaa	tgaaaattcca	gggtgggtcg	aattgacggc	agactataac	gattgactgt	480
aagaaaattgt	actctgatac	gctgggggtg	agtgggttta	atgtrgaatg	tagcgggggt	540
gtggctcaata	ctgcagcgac	taaatctgat	ttggcagcag	ctcaactctt	ggctccaggt	600
actgctgagtg	ctaattggatc	agtttaacct	actgctggcg	caggcctgaa	acacattaca	660
gctgcagatg	taattgcagag	tttggcttaac	aacgcgaacg	ttaatggccc	aattgcaaat	720
gggtttggatg	cgccaacagc	tacagattat	actatacaac	gcgtgacacg	cgattttaca	780
tatatgtgcaa	ctatttgcagc	tggttacaaac	tcctgggtata	gtaaacagtc	tcagtttaca	840
tccttctcga	caccaaaagc	gggcgataact	gctaacttaa	acgttcaaac	tggtttctacg	900
tcgaattggc	ttgtattggc	tagcgacggt	aaaaattaccg	cgaagagatg	ttcagataca	960
ttatttgacg	tagatggtaa	ctcactcaaa	acaaatgctg	cgagtcgctt	atctacataca	1020
cttgatgcac	tgactaaaaa	ctggcctaca	acaggcacac	cgagtcgctt	tactactctt	1080
atcacaaactg	aagatgaaac	aaacttcaact	ctggctggcg	gtactagctg	tactactctt	1140
gggtgcaatca	ctgtagcaaa	tgcaagaatg	agtgtctgag	ctcttcaact	ggcacaactag	1200
tcacacagat	tcacagttga	tggtggagct	actggtgac	gcgcagcgca	tattaaaggt	1260
gatagttaag	gtatagtaca	acacacacac	ggtacaggtt	ttgaagacgc	ttacacccaa	1320
gctgagtgggt	cactgactac	cgataataca	accaaactgt	ttttgcaaaa	agacggaaact	1380
gtgaccgaatg	gttccagtaa	agcagctctat	gtttcagcgg	atggtataatt	tactactgac	1440
gctgcaaaa	aagttgcac	caccgcgcat	ctctgaaag	ctctggacga	agcagtcagc	1500
tcactgcgaca	aaatccgttc	ttccctcgtg	gcggtgcgaa	accgcttgga	ttccgcagtc	1560
acccaactga	acacaccac	tactaacctg	tcgtgaagcg	gttcccgatg	tcaggacagc	1620
gactatgcga	cgaaggtgtc	caatatgtcg	aaagcgaga	tcattccagca	ggccggtaac	1680
tcctgtctgg	caaaagctaa	ccaggtacgc	cagcaggttc	tgtctctgct	gcagggttaa	1740

<210> 28

<211> 1233

<212> DNA

<213> Escherichia coli

<400> 28

acaaaaaac	agatctgcgt	gtcgacttct	atcgagcgcc	ttctctctg	ttctgcgcat	60
aacagcgcta	aagatgacgc	tgccggggcag	gcgatttgcta	accgcttacc	ttctaacatc	120
aaaggtctga	ctcaggccgc	acgttaacgcc	aacgacggta	ttctctctgc	gcagaccact	180
gaagcgcgac	tgctgtaaat	caacaacaa	ttgcagctgtg	ttcgtgagct	cagcgtttac	240
gccactagcg	gtactaactc	tgattctgac	ctgtcttcaa	ttcagggcga	aatacaattc	300
cgctctcagtg	aaattgaccg	cgatcccggt	cagactcaagt	tcacgcggct	gaagctactg	360
gcataaagata	acacccatgaa	gattcaggtt	gggtgcgaac	atggtcagac	tataccatc	420
gacctgcgaa	aaatcgactc	ttctactctt	ggtttgaacg	gtttctccgt	ttccaaaaat	480
gctctcgaaa	ctagcgaagc	gatactcag	ttgccgaacg	gtgcgaatgc	accaaactgc	540

gtgaagatg	atgcgtctgt	tctgaccgat	cttaacatta	ctgatgcttc	cgctgtttcg	600
ctgcacaa	taactaaagg	tgggtgcga	acgtctactt	atgtgtttca	gatatgcgat	660
aagagctatg	cagcatctgt	tgatgcccga	ggtaacagta	aactgaataa	agccgacgta	720
acataataat	acgcagcaaa	tgggtgttacg	aatgccacc	agattggtag	tctgtgttcag	780
gttgggtctg	atgcacaa	tgatgcagtt	ggttttgtta	ccgtgacagg	gaaaaactat	840
gttgcctaag	actcattagt	caatgcta	ggcgtctgctg	gcgtgcagc	aactagagtt	900
acaatgatg	gtgatggtag	ccttggagct	aaccagctta	aaattgaact	tagccaaat	960
gggtgcactg	ctgcacatc	agagttcgct	gggtgcttca	ccaacgactc	acrgactctg	1020
ctggacaaat	ctatcgcatc	tgttgataaa	tctcggttct	ctttggggcg	ggtcagagac	1080
cgcttgagct	ccgctgtaac	caactggaac	aacacacaa	ccaacctgtc	tgaagcgacg	1140
tcccgtatct	aggacgcga	ctatgcgacc	gaagtgtcca	acatgtcgaa	agcgcagatc	1200
atccagcagg	caggtaactc	cgtgtgctcc	aaa			1233

<210> 29  
 <211> 1713  
 <212> DNA  
 <213> Escherichia coli

<400> 29						
atggcacaag	tcattaatat	caacagcctc	tgcgtgatca	ctcaaaataa	tatcaacaag	60
aaccagctctg	cgctgtcgag	ttctatcgag	cgctctgtctt	ctggctttgag	tattaaacgc	120
gcgaaggatg	acgcgcgagg	tcaggcgatt	gctaaccgctt	ttacttctaa	ctattaacagg	180
ctgactcagg	ctgcacgttaa	cgccaacgac	ggatattctg	ttgcacagac	cactgaaggc	240
gcgctgtccg	aaatcaacaa	caacttacag	cgatattctg	aactgacggt	tcaggcgacg	300
acgggaactat	actccacttc	tgaactggac	tccattcagg	acgaaatcaa	atcccgctct	360
gatgaaattg	acgcggtatc	cggtccaaac	cagttcaagg	gcgtgaacgt	actgtcaaaa	420
gatggctcga	tgaaaattca	ggctggcgca	aatgatggg	aaacactcac	gattgatctg	480
aaaaagatgc	actctctctc	atgtgaagctg	acacgcttca	atgttaacgg	taaaaggcgt	540
gttgataatg	ctaaaagcac	tgaagcagat	ctgacgcgctg	cgcgctttca	ccaaggtgca	600
gtcgtcagtg	gcaacagcac	ctggactaaa	tctactgtta	ctacotttaa	tgcagcaaca	660
gtacocgacg	tgctggcaag	cgctagcggc	ggcagcacta	ttagcggtta	tcccggtaca	720
aacaattgat	taggcgtagc	ggcttctact	gcataacct	acaacgcaac	cagcaagctt	780
tattcatttg	acgcaacgcg	acttaccact	ggcgatggt	ctggggccac	actcaagtt	840
gctgatgtgc	tgaagcgtta	tgcagcaaac	ggtgataata	cggtccagat	ctccactcgg	900
ggagcgctc	aggcgtttaa	aattgcgcg	gatggcaacc	tgactgaagt	caatggtagt	960
gctttatata	acggttctga	cggaacactg	actaaaaacc	agccgcgcgc	tcagatgtcg	1020
gcaacgcttg	acgggtattt	caacgggtgc	aattgtaagt	cagcagttga	tcggaagatt	1080
acattcgcca	ggcgcatgac	cggtgatctc	acccagctta	gcaaaaaagt	ggatataaag	1140
ggcgcaacgg	tatccgcgca	agatattgac	actgcgttaa	ctggcgagcg	ttataacgta	1200
gctaaccgcg	caacagcttt	tgcagttgac	gctgggtggg	cagtaacgcg	tactaacggt	1260
ggcgctacgg	taaatattgg	tgtgtatggt	gaactgacga	ctgcgaccaa	caagactgtc	1320
acagaaactt	atacgaatt	tgttaacgcg	aatattctcg	atgatgacg	cgcggtctgt	1380
tacaacacgg	ctgacggttc	tctgaccact	gaagctactg	gtaaatccga	agtgacacac	1440
gatccgctga	aagcgtggga	cgatgctatc	gcactcgtag	acaaattccg	ctctcccttc	1500
gggtcggtgc	agaaacgctc	ggattccgca	gtcccaaac	tgaacaaac	cactcaaac	1560
ctgtctgaag	cgcatctccg	cattcaggac	gcgcagtag	cgacggaagt	gttattatgt	1620
tgcgaagcgc	agatcatcca	gcaggccggt	aactccgtgc	tggcaaaagc	caaccaggta	1680
cgcgacgacg	ttctgtctct	gctgcagggt	taa			1713

<210> 30  
 <211> 1668  
 <212> DNA  
 <213> Escherichia coli

<400> 30						
atggcacaag	tcattaatat	caacagcctc	tgcgtgatca	ctcaaaataa	tatcaacaag	60
aaccagctctg	cgctgtcgag	ttctatcgag	cgctctgtctt	ctggctttgag	tattaaacgc	120
gctaaggatg	acgcgcgagg	tcaggcgatt	gctaaccgctt	ttacttctaa	ctattaacagg	180
ctgactcagg	ctgcacgttaa	cgccaacgac	ggatattctg	ttgcgacagc	cactgaaggc	240
gcgctgtccg	aaatcaacaa	caacttacag	cgatattctg	aactgacggt	tcaggctctt	300
acgggacta	actccgattc	ggatctggac	tccattcagg	acgaaatcaa	atcccgctct	360
gacgaaattg	acgcggtatc	tgcccgagac	cagttcaacg	gcgtgaacgt	actgtcgaaa	420
gacggtttca	tgaaaattca	ggttgggtgc	aatgacggcg	agactatcac	tattgatctg	480
aagaaaattg	actcagatc	gctggggcgtg	agtgggttta	atgtgaatgg	tgctggggctg	540
gtgtcaata	ctgcagcgac	taaaagtgat	tgtgctgctg	catcagtttc	agctgcgcta	600
ggtaattgat	acactgtctc	tgtctggcgtg	tgcgaatcaa	ctgtctgctg	tgattatgtg	660
agttctcacg	atggtgcgac	agtaactcgc	gctggtgtaa	gcaatggttt	tgtctcaggg	720

gcaactggag	atgcttataa	attcaatcaa	gcaaacaca	cttttactta	caataccacc	780
lcaacagcgg	cagaactcca	attctactcc	acgcctcaag	cgggggatac	cggaactttc	840
tcogtggaaa	ttgtgggcac	caagcaggat	gtgtgtctgg	ctagtgtatg	caaaatcaca	900
gcgaaaagacg	ggcttaaaact	ttatattgac	accacagggg	ttagctttca	aaaagggtga	960
ggtaactttag	aagaagctac	ctctcaatggc	ttagctttca	accactctgg	tccagctgct	1020
gctgtacaat	ctactattac	tactcgggat	ggaactctca	tagttctagc	agggttctggc	1080
gactttggaa	caacaaaaac	tgtctggggc	attaaatgta	caggagcagt	gattcattgct	1140
gatgcacttc	ttctcccgac	taaagcgact	gggtttactt	ctggcactta	taccgttaggt	1200
acagatggag	tgttaaaatc	tgggtggcaat	gacgtttata	acaaagctga	cggggacggg	1260
ttaactactgt	acaataccac	aaaatatatat	ttaacagatg	acgggtctgt	aaactaatggt	1320
tctggtaaaag	ctgtgtatgc	tgaatgcaaca	ggaaaactaa	ctactgacgc	tgaactataa	1380
gcgcgaaacca	cccgcgatcc	cctgaaaagct	ctggaacgaag	cgatcagctc	catcgacaaa	1440
tctcgctcttt	cctctgggtg	gggtgcgaaac	gtctctggat	cgccggctca	caactcgtaac	1500
aaacacaccta	ccaacctgtc	cgaagcgacg	tccgtatttc	aggacgcccga	ctatcgaccac	1560
gaagtgtcca	acatgtcgaa	agcgagatgc	atccagcagg	cgggttaact	cgtgtgtggca	1620
aaagctaaccc	aggtaccgca	gacgggtctg	tctctgctgc	agggttaa		1680

<210> 31  
 <211> 1713  
 <212> DNA  
 <213> Escherichia coli

<400> 31						
atggcacaag	tcattaatcc	caacagcctc	tcgctgatca	ctcaaaaata	tatcaacaag	60
aaccagctctg	cgtctgcgag	ttctatccag	cgtctgtctt	ctggctctgg	tattaaacgc	120
gcgaaagatg	acgcgcgggg	tcaggcgatt	gctaaccggt	ttactttcaa	cattaaaggc	180
ctgactcagg	tgcacgttaa	cgcacaacgac	ggattttcgg	tgcgcagac	acccgaaagc	240
gcgctgtccg	aaatcaacaa	caacttacag	cgtatccgtg	aactgacggt	tcaggccaact	300
accgggtatc	actccgattc	tgaactggac	tccatccagg	acgaaatcaa	attcgtctgt	360
gatgaaattg	accgcgctat	tggctcagacc	cagttcgaatg	gcgtgaaatg	gtgtcccaaa	420
gcagggttcaa	tgaaaattca	gggtgggcga	aatgatgggt	aaacatcac	gattgacactg	480
aaaaaaatcg	actctctcac	actgaagctg	accagcttca	acgtcaacgg	taaaagggcg	540
gttgataatg	caaaaagccac	tgaagcagat	ctgacccgtg	cgggcttctc	ccaaagtggc	600
gttgctcagtg	gcaatagcac	ctggactaaa	ttactctgta	ctacctttaa	tgacgcaaca	660
gtctaccgatg	ctgtggctag	cgttagtgag	ggcagcacta	tacgctgta	tgctgggaca	720
aacacagggt	taggcgtagc	ggcttctact	gcatacacta	acaaagccac	cagcaagctct	780
tattctatttg	acgcaacgcg	aactactaat	ggatgatgta	ctgcccggct	aaactaaagt	840
gctgatgttc	tgaagcctta	tcgacgaaac	ggcagataaa	cggctcagat	ctccactggt	900
ggtagcgtct	aggaagttaa	aattgccagc	gatggtaacc	tgaacgatac	taatggcgat	960
cttttataca	tgggtgctga	cgttaacctg	acgaaaaaac	cccgccggcg	cccgccggcg	1020
gcctaagcttg	accgtatttt	caacgggtgcg	aatggtcatg	atgcagttga	tgccgaagat	1080
accttcggca	cggcgatgac	cgttgacttc	accaggttta	gcacaatgt	ggatattaa	1140
ggcgccagcg	tattccgcga	agatatgaac	actgcgttaa	ccggtccagg	ttatacctga	1200
gctaaccgcg	caacgtctta	tgaactgtgc	gctgatgggt	cagtaactgc	caactacggt	1260
ggagcgaccc	tataatattg	tgtcgaggtg	gaactgacga	ctgcggccaa	caagactgct	1320
acagaaacct	atacgaatt	tgttaacggc	aatattctgg	atgatgacg	cgcgctctgt	1380
tataaagcgg	ctgacggctc	tctgaccact	gaagctacag	gtaaatctga	acgacgcacg	1440
gctacgctga	aagcgttggg	cgtatgctatc	gcactccgtg	acaaaattcg	ttcttctcgt	1500
gggtccgctgc	agaacgtctc	ggatttccga	gtcacaacac	tgaacaacac	cactacacac	1560
ctgtctcgaa	cgcagctccg	tattcaagac	gcgcgactag	gtaacgaagt	gtcccaactg	1620
tgcgaagcgc	agattattca	caggcgaggt	aactccgtgc	tggcaaaaag	taaccaggta	1680
ccgcgacgag	ttctgtctct	gctgcagggt	taa			1713

<210> 32  
 <211> 1188  
 <212> DNA  
 <213> Escherichia coli

<400> 32						
aacaaaaaac	agtctgcgct	gtcgacttct	atcgagcgcc	tctctctctg	tctgcgcaat	60
aacagcgcta	aaagtgcgac	tgcgggcgac	gcgattgcta	accgcttca	ttctaactac	120
aaaggtctga	ctcaggccgc	acgttaacgcc	aacgacggtg	tctctctggc	tgacacactc	180
gaagcgccac	tgtctgaaat	caacaacaac	ttgcagcgtg	tgcgtgagtt	gactgtctag	240
gcgacgacgc	ggactaacct	tgtattctgac	ctgtcttcta	ttcaggacga	atacaaatct	300
cgtctgggatg	aaattgacgc	tgtttccggt	cagacccagt	tcaacggcgt	gaagctgtgt	360
gtcttcgaa	gtctctatgc	gattcagag	gcgcgactag	atgggcagac	cattcaactc	420
gacctgcaga	aaatcgactc	ttctactctg	ggcctggggc	gcttctccgt	attcaacaat	480

gcactgaaac	tgagcgattc	tatcactcag	gttggtgcga	gtggttctact	ggcagatgtg	540
aaactgagct	ctggtgcctc	ggctctgggt	gtagacgcga	gcactctgac	tctgcacaac	600
gtacagacc	cagctgggoc	agcaacagct	aactatgttg	tctctctcgg	tctgcacaac	660
tactcagtat	ctgttgaaga	tagctccggg	acagttacgc	tgaaccaccac	tgatataagg	720
tataccagata	cgcctaattg	cgttactacc	ggttccatga	ctggcaagta	cgttcaagtt	780
ggagctgatg	cattgggtgc	tgtctgtagg	tatgtccacc	tacaggggaca	aaacttcaaa	840
ctgatgtctg	gcgcgctggg	taactccaag	aattgctcgt	gtagtccaga	tgttactctt	900
gcaattggcg	atatgtctaa	taagcggaat	gctaaccatt	acactgggaac	ctctctctcga	960
gtccactcgg	ctatgtctgga	caaaactatc	gcattctgtg	ataaaattccg	ttctctctcta	1020
ggggcggtgc	agaaacgctc	gagctctgct	gtaaccacaac	tgaaccacaac	cactaccacac	1080
ctgtccgaag	cgcagctccg	tattcaggac	gcgcactatg	agcccggaag	gtccaaacatg	1140
tcgaaagcgc	agatcatcca	gcaggcggtg	aactccgtgc	tgtctaaa		1188

<210> 33  
 <211> 1638  
 <212> DNA  
 <213> Escherichia coli

<400> 33						
atggcacaag	tcattaatac	caacagcctc	tcgctgatca	ctcaaaaata	tatcaacaag	60
aaacagctctg	cgcgtgcgag	ttctatcgag	cgtctcctct	ctggctctgg	tattcaacagc	120
gcgaagcgatg	acgcgcgcgg	tcaggcgatt	gctaaccgtt	ttacttctaa	cttaaaggc	180
ctgctactcag	ctgcacgtaa	gcctaactgac	ggatattctg	ttgcacagac	cactgaaggc	240
gcgctgtccg	aaatcaacaa	caactctacg	cgtatctcgg	aactgacggt	tcaggtctct	300
acgggagcata	actctgattc	ggatctggac	ttcatctcag	acgaaatcaa	atccgctctc	360
gacgaatttg	acgcgctatc	cggctcagacc	cagttcaacg	gcgtgaacgt	ctcgtgcaaaa	420
gacggctcga	tgaaaaattca	ggttgggtgc	aacgacggcc	agactatcac	tattgatctg	480
aagaaaatttg	actctgatac	gctggggctg	agtggtttta	acgtaaatgg	tagcgcagat	540
aaggcgaaagt	tgcgcgcgac	agctgacgga	atggctaaag	acggatatac	caaaagggta	600
acttcatctg	acggcagcac	tgcataatac	aaaaactcac	caaatctctg	agcaaaagac	660
ttctgatattc	tgcgcgcgct	taagactcgc	gataaaaata	cgcgaacagg	tgcaaaatagc	720
cttgcgtgata	atgcgcacac	gacaaactat	acttataatt	caaccagcaa	taccttctcc	780
tatacggctg	acgggtgaaa	ccaaacgaat	gctgcagcaa	atctcatacc	tgacgtcagg	840
aaaaacgacg	ctgcatcagt	tactattggg	gggacgacac	agaaatgaaa	tatttgattg	900
tcgggcaata	ttacttcaag	tgtatggcgt	caactctatc	tggaattcaac	aggtaacctg	960
actaaaaaac	aggccggcaa	ccgaaaaaaa	gcaaccgttt	ctgggcttct	gggaaatacg	1020
gatgcgaaag	gtactcgtg	taaaacacac	atcaagacag	aggtcgtgtg	aaaggttaca	1080
gctgaaggta	atacaggtac	tgtaaaaatt	gaaggtgcta	ctgcttcagc	atctgcattt	1140
acggcgatgt	ctatttcgcg	caacaacggg	gggaataact	atgctgtctg	cgcaataaat	1200
actcaacaatg	gtttctctgc	gggggatgac	ttaaaccagc	atgctcaaac	tggttcaaac	1260
tactactcgt	aagccgatgg	cacggtcacg	aatagcgcaag	gcaagaataat	ctataaagac	1320
ctgtgatggg	tctacagcac	agagaataaa	acatcgaaag	cgtccgatcc	attgctcgtc	1380
gttcgacgac	caactagctc	catcgacaaa	ttccgttcat	ctctgggtgc	tatccagaaac	1440
cgtctggattc	cgcgggtcac	caactctgaac	aaacccacta	ccaaactgtg	cgaagcgcag	1500
ttccgtattc	aggaacgcga	ctatgcgacc	gaaggtctca	acatgtcgaa	agcgcagact	1560
atccagcagg	cggtaactc	cgtgctggca	aaagctaacc	aggtaccgca	gcaggttctg	1620
ttctctgctg	agggctaaa					1638

<210> 34  
 <211> 2145  
 <212> DNA  
 <213> Escherichia coli

<400> 34						
aacaataatc	agtcttctct	gagctccgcc	attgaacgtc	ttcttctctg	cctgcgtatt	60
aaacagtgcta	aagatgaagc	agcaggtcag	gcgatgtcta	accgttttac	agcaaatatt	120
aaagcgtctga	ctcaggcttc	ccgtaaacgc	aatgatggta	ttctgttgcg	gcagcaacct	180
gaaggtgcgc	tgaattgaat	taacaacacac	ctgcagcgtg	tacgtgaact	gactgttccag	240
gcaactaacg	gtacttaact	tgacagcgat	ctttctctga	tcagggctga	aaacttcaaa	300
cgtctgggaag	aaattgaacg	tgtatctcag	caaaactcag	ttaacggcgt	gaaagctct	360
gctgaaaaata	tgaataatgaa	aattcaggtt	ggtgctaat	atggtgaaac	ctactactat	420
aattctggcga	caattgatgc	gaaaactctc	ggcctggacg	gttttaatat	cgatggcgcg	480
cagaaagcaca	cctgcagctga	ctgattttct	aaatttaagc	cgacaggtac	tgaataactat	540
gatgttggcg	gtgatgctta	tactgttaac	gtagatagcg	gagctgggta	atgactccaa	600
cttatgtgata	gtgttttatg	ttcagataat	gcccgatgac	tttgtcatgc	agctccacgc	660
attttgagaa	cgacagcgac	ttccgtccca	gccgtgcaag	gtgctgcctc	agattcaaggt	720

tatgcgcgtc	aattcgctgc	gatatcgctc	tgtcgattac	gtgcagcttt	cocttcaggc	780
gggattcata	cagcgccagc	ccatccgctc	tccatcatc	cacgtcaaa	ggcgacagca	840
gggtccatag	agcccccagc	gtcgccatag	tgcgctcacc	gaatacgtgc	ggcaaacacg	900
tcttcgcggag	ctctgcatac	gcgtataaac	gccagcgctg	gcgcgattta	ggcccgacat	960
agtcoccatg	agtcgccatt	tccgcgcaga	cgatgacgtc	actgcgccgc	tgatgcgcgc	1020
aggttacccga	ctgctggcctg	agttttttga	gtgacgtata	actggtgtga	ggcaaacacg	1080
cataatgcggg	gcagttgccc	ggcatccaac	gccattcatg	gccatataca	tgattttctg	1140
gtgcgtaccg	gggtgagaag	cgtgtgaagt	gaactcgagt	tgccatgttt	tcaggcagtg	1200
agagcagaga	gagcgctgat	gtccggcggt	gcttttgccg	ttacgcacca	ccccgtcagt	1260
agctgaacag	gagggcacgc	tgatagaaac	agaagccact	ggagcacctc	aaaaaacaca	1320
tcatacacta	aactcagtaa	ttggcagcat	taccgcgagg	ctgttaaaga	tactacaggg	1380
aatgatattt	ttgttagtgc	agcagatggt	tccactgaca	ctaaactctg	cacaaacata	1440
gctggtacag	gggattgatgc	tacagcactc	gcagcagcgg	ctaagaataa	agcacagatt	1500
gataaattca	cggttaattgg	agttgaattc	acaacacaaa	ctgcagcgga	tgcccaatgg	1560
aatgggtgat	attctgcaga	aattgatggg	agtcagtgga	catttactgt	gacagatgct	1620
gacaaaaaag	ctctcttgat	tacgagtgag	acagtttaca	aaaatagcgc	tgccctttat	1680
acgacaccca	aagttgataa	caaggctggc	acactttcgc	atcttgatct	caatgcagct	1740
aagaaaaacg	gaagcacggt	agttgttaac	gggtcaactt	acgatgttag	tcgatcaggt	1800
aaaaacgata	cggagactgc	ttctggtaac	aataaagtca	tgatctctag	caaatcagaa	1860
gggtggtagcc	cgattctcgt	aaacgaagat	gcagcaaaat	cggtgcaact	taccaccaac	1920
ccgctcgaaa	ctatgcacaa	agcatggctc	aaagttagca	atctgcgctc	tgacctcggg	1980
cgagtacaaa	acggtttcga	ctctgctatc	accaactctg	gcaacacgct	aaacacactg	2040
tctcttgccc	gtagccgtat	cgaagatgct	gactacgga	ccgaagtgtc	taacatgtct	2100
cgtagcgaga	tcttgcaaca	agcgggtacc	tctgtctcgg	cgcag		2145

<210> 35  
 <211> 1587  
 <212> DNA  
 <213> Escherichia coli

<400> 35	aacaagaacc	agtcctgcgt	gtcaggttct	atcgagcgtc	tgtctcttgg	cttgcgtatt	60
	aacagcgcga	aggatgacgc	gcgaggtcag	gcgattgcga	accgttttacc	ttctaacatt	120
	aaaggcctga	ctcaggtctgc	acgtaacgcc	aaagcagctg	ttctgtttgc	gcgacacacc	180
	gaagcgcgcg	gtgcgaatat	caacaaacac	ttacgtagct	tcgcgtgaact	gaacgcttga	240
	gcacacacgc	gtacccaactc	ccagtctgac	ctggactcta	tcagagacga	aattaaactc	300
	cgtctggcag	aaattgacgc	cgtatccggt	cagacccagc	tcaacgcggt	gaacgtactg	360
	gcacaaagac	gttccatgaa	aattccaggt	ggcgcgaaag	atggcgacag	catcactatc	420
	gacctgaaga	agattgactc	ttctacaggt	aaactgactg	gttttaacgt	gaatggcaaa	480
	ggcgctgtgg	attcaaatgg	actagcagca	gttaaagatg	gcagcacact	aaattacacc	540
	gcggcaactg	cagtaaacgt	actagcagca	gttaaagatg	gcagcacact	aaattacacc	600
	ggctactgga	atggttttag	gattgctgca	acaagtgtct	atcacattca	cgatagcact	660
	aaatccctata	cccttgattc	tacgggggct	gcagttagctg	gtgcgcgctg	gaacctgcaa	720
	gggtactttg	gtacagatgc	gaatactgca	aaaatcacca	tcgatgtgtc	tgctcaagaa	780
	ttcaaacatgc	ctaaagatgg	gaaaatttact	gatactgtag	gttaaagctt	atatatcgat	840
	gtcctactgt	gttgacttaa	gaacggctct	gatactttaa	ctcaggaac	attgaatgat	900
	gtccactcgt	gtgtctaatt	agttgatgat	acaaggattg	acttcgatag	cggatgtctt	960
	gtccacccctg	ataagtgtaa	cagcactgta	gatacactgt	gcgcacttat	ttcagccgct	1020
	gcaatgacta	ataagttgac	aggttaagcc	tataccgtag	taaatgggtg	agaaattctt	1080
	gctgtgacta	ctaataacac	agtaaaaacg	actgctgatg	ctaaaaatgt	tatgtttgat	1140
	gctgtgacta	ctaataacac	tgatgacaaa	gcaactgtta	gcagaaacta	tcagtgaatt	1200
	gcgaatggca	atactatgta	tgataaaggc	gctgtctgtt	atgcggcgcc	ggatgtgtct	1260
	ctgactacag	aaactacaga	taaatcagaa	gctacagcta	accgcctggc	cgctctggac	1320
	gcgcgaatca	gcagatgcga	caaatccgtg	tcactccctg	gtgctatcca	gaacgctctg	1380
	gattctgcag	tcaccaactc	gaacaaaccc	actaccaact	tgctgtgaag	cgaagtccgtg	1440
	attcaggagc	ccgactatgc	gaccgaagtg	tccaatatgt	cgaagcgcca	gatcatccag	1500
	caggcaggca	actccgtgct	ggcaaaaa				1560

<210> 36  
 <211> 1245  
 <212> DNA  
 <213> Escherichia coli

<400> 36	aacaaaaaac	agtcctgcgt	gtcaggttct	atcgagcgtc	tctctcttgg	cttgcgcatt	60
	aacagcgcga	aagatgacgc	gtcggggccag	gcgattgcga	accgttttacc	ttctaacatt	120

aaaggtctga	ctcaggccgc	acgtaacgpc	aacgacggta	tctctctggc	gcagaccact	180
gaaggccgac	tgcttgaat	caacaacaac	ttgcagcggt	ttcgtaact	gacggtctag	240
gcactacccg	gtactaacct	tgattctgac	ctgtcttcaa	tcacaggacga	aatcaaatcc	300
cgctctcgatg	aaatgacccg	cgatccgggt	cagactcagt	tcaacggcgt	aacgtactgt	360
gcataagatg	agatgacgaa	aattcagggt	gggtcaaatg	atggtcagac	aatcagatt	420
gatttgcaga	agattgatct	tcttacttta	gggttaaatg	gttttctgt	ttccaaaaat	480
cgagatctcg	ttggtgatgc	tattactcaa	ttgctctggg	agaaggcagc	cgatgcacca	540
gtaacctcata	agtttgatga	ttcagtaaaa	actgatttaa	aactgacgga	tgcttcaagg	600
ttaaagctcgc	ataacctcaa	agatgaaaa	ggataattaa	ctaacccagta	tgtgtgacag	660
aatggcgga	aattcttcgc	tgctacagtc	gctgcgaatg	gtatgtttac	gtcgaacaaa	720
gcataatgtaa	ctcacagcga	tgctgcgaac	ggatattgata	ccgcaacgca	gtcaggccag	780
ttagttcagg	ttggtgcaga	tcttaccggg	agcgcacaa	catctgtgtc	tgcttcaagg	840
aaaagctctg	tcggtgatga	cgccgtctgc	aagaataaca	ctgtgtgatgc	tacgctact	900
caacggggaa	catctgggac	aacgggtctc	gcagcgctcaa	ttcatctgag	tcggcgcaaa	960
aactctgtgag	acgtgatgt	acgggtctcc	actgaattca	caggtgtctc	aaccaacgat	1020
ccactgactc	tgctggacaa	agctatcgca	tctgttgata	aattccgttc	ttctttgggg	1080
gcggtacaga	acgtcttgag	ctccgctgta	accaaactga	acaaacccac	caaccaacctg	1140
ctctgaagcgc	agtcctgata	tcgagcgctc	gactatgcga	cggaaagtgc	caacatgtgc	1200
aaagcgcaga	ttatccagca	ggcaggtaac	tcctgtgtgt	ccaaa		1245

<210> 37  
 <211> 1185  
 <212> DNA  
 <213> Escherichia coli

aaacaaaaac	agtcctcgct	gtcgactctc	atcgagcgcc	tctctctgg	tctgcgcatt	60
aacagcgcta	aagatgacgc	tgccggccag	gcgattgcta	accgcttcac	ttctaaccac	120
aaaggtctga	ctcaggctgc	acgtataacgc	aattgacggt	ttctcttagc	acagacagcg	180
gaaggcgccg	tgatcagagat	taacaacaac	ttgcagcggt	tgctggaagt	gacgctgcag	240
gcaaccactg	gtaccaactc	tgattccgat	ctctctctca	ttcaggatga	aattaaatct	300
cgctctggatg	aaattgacgc	cgctctctgt	cagacccagc	ttaacggcgt	gaacgtactg	360
gctaaaaaacg	gtctctatgg	aattccaggt	ggcggaagc	atggccagac	tatctctatc	420
acgctgcaga	aaatagaccc	ttctactctg	ggctcgagtg	gcttctctgt	ttctcagaaa	480
tcctctgaac	tgacgcatcc	tatcactacg	ctcggcaata	ctactgtctg	atcgaagaac	540
gtggagctga	ggcgagtagc	aactaaaactg	ggctggaatg	caagcacctc	gagcctgcac	600
gaagttcagg	actctgctgg	tgacggtagc	gtgactcttc	ttgtttcttc	tgccagcgac	660
aaactatgctg	tgctctgaga	cgcggtctct	ggtagcagta	acctgaacac	actgacgact	720
acctatgatg	acgtactact	tggtgttact	ggcgagcact	agaaacggta	gctgactcaa	780
gtaacttctg	acgccaacgg	tgacgtctgt	ggttacgtaa	ccattcaggg	taaaaactat	840
caggctgggtg	gcagccgggtg	tgacgtctctg	gcgaacagcg	gtgtgtcagc	ttcaactaca	900
gctgttgata	ccgttactct	gcaactgagc	ggtagctgtg	caactacgta	tgataaagg	960
actgcgaact	agaaacccact	ggcactattg	gacaaagcta	tcgtctctgt	tgataaattc	1020
cggttctctc	tggtgtcggt	acagaactct	ctgagctctg	ctgtaaccac	ctgtaaatat	1080
acacacacta	acctgtctga	agcgcagctc	cgatttcagg	atgcgcagta	tgcgaccgaa	1140
gtgtcaaat	tgcttaaacg	gcagatcgtt	cagcaggccg	gtaac		1185

<210> 38  
 <211> 1383  
 <212> DNA  
 <213> Escherichia coli

aaacaaatctc	agtctctctc	tagctctgct	attgagcgct	tgctctctgg	tctgcgtatt	60
aacagcgcaa	aagacgatgc	agcaggtcag	gcgattgcta	accggtttac	ggcaaatatt	120
aaaggtctga	cccaggcttc	ccgtaacgca	aatgatggta	ttctctgttc	gcagaccact	180
gaaggtgcgc	gtactgaat	taacaacaac	ctgcagcgta	ttcgtagaact	tctctgtcag	240
gcaactacgc	tgactaacct	tgacagcgat	ctttctctca	ttcagcgctga	aattactcca	300
cgctctggaa	aaattgacgc	tgattctgag	caaaactcagt	ttaacggcgt	gaagactcct	360
gctggaataa	atgaaatgaa	aattcaggtt	gggtgcaatg	atgggtgaaa	catcactact	420
aatctgggcaa	aaattgatgc	gaaaactctc	ggcctggagc	gttttaatat	cgatggcgcg	480
cagaaagcaa	caggcagtag	cttgatttct	aaatttaaa	gcacaggtga	tgtataattat	540
gatgttgccg	gtaaaaacta	taccgtgaat	gtggagagcg	gcgcgggttaa	gaatgatgct	600
aataaagatg	ttttgttaag	cgacagctgat	ggatcgctga	cgaccagtag	tgataactaaa	660
gtatccgggt	aaagtattga	tgcaacagaa	ctagcgaaac	ttgcaataaa	attagctgac	720
aaaggtctcca	tgtaatacaa	ggcatttaca	tttacttaaca	acactggcgc	agagcttgat	780
gctaagtggta	aaggtgtttt	gaccgcaaat	attgatgtgtc	aagatgttca	atttactatt	840

gacagtaaatg	caccocacggg	tgcggcgcgca	acaataacta	cagacacagc	tgtttaca	900
aacagctggcg	gccagcttcac	cactacaaaa	gtggaaaaata	aagccgcaac	actctctgat	960
ctggatctcta	atgcagcccaa	gaaaacaggt	agcactttag	ttgtaaatg	cgccactctac	1020
aatgtcagcgg	cagatggtaa	aacggtaact	gatactactc	ctgttgcccc	taagtgatg	1080
tatcttcagca	aatcagaagg	tgttagcccg	atctctggtaa	acgaagatgc	acgaataatg	1140
ttgcaatctta	ccaccacccc	gctgaaact	atcgacaagg	catggctaa	agtgacaaat	1200
ctgcgttctta	actccgggtg	agtcacaaa	cggtttgact	ctgccatcac	caactctggc	1260
aacacogtaa	acaactctgc	tctgcccgt	agccgtatcg	aagatgctga	ctacgcgacc	1320
gaagtgtctta	cactgtctgc	tgccgagatc	ctgcaacaag	cggttacctc	tggtctggcg	1380
cag						1383

<210> 39  
 <211> 1680  
 <212> DNA  
 <213> Escherichia coli

<400> 39						
atggcagcag	tcattaatc	caacagcctc	tcgctgatca	ctcaaaaata	tatcaacaag	60
aaccagctctg	cgctgcgag	tctatctcag	cgctctgtct	ctgctgttgc	tattacaagc	120
gcgaaagatg	agcccgccag	tcaggcgatt	gctaaccggt	tcactcttaa	cattaaaggc	180
ctgactcagg	ctgcagctaa	cgccaacgcag	ggatattctg	ttgcacagac	caccgacgct	240
gcgctgtccg	aaatcaacaa	caactttacag	cgatccggt	aactgacggt	tcaggctctt	300
acgggagacta	actctgattc	ggatctggag	tccatccagg	acgaaataca	atcccgctct	360
gacgaaatttg	acgcgctatc	cgccagcagc	cagttcaaac	ggctgaacgt	ctggcgcaaa	420
gacgggttcaa	tgaaaattca	ggttgggtgc	aatgacgggc	agctatctac	tattgatctg	480
aagaaaaattg	actctgatac	tctgggtttg	agtggattta	atgtgaatgt	caaaagggtc	540
gtggctaaag	caaaaagcag	cgaagcagat	ttacgggggg	ctggtttctc	tcaaggagcg	600
gtggatataca	acgggaaatg	tacttggaac	aaatcaacac	ccaccaatca	ctcagctgca	660
acaaactgctg	acttgttatt	gaccattaa	gatggctcta	ctgttatcata	tgacgggaca	720
gacacccggt	taggggtcgc	agcagcagga	aattatactt	atgatgcgaa	cagtaaatct	780
tattctcttca	atggcaatgg	tctgacgggc	gcataaaaat	caactgcact	caaaaggtac	840
ttgggggacag	gtgctaacac	cgctaaaaat	tctatcggtg	gtacagagca	ggaagtgaa	900
attgcccaga	atggcactat	tacagatacg	tatgggtgat	cgctctatct	ggatattaac	960
ggcaactcgca	ctaagaacta	tgcgggttca	caactctgca	caacgctgga	taacgtatta	1020
ctgtcccgaa	cgttataatg	cactataca	tttgatagcg	gtatgacggt	tgattcaact	1080
cgagctactg	ggcggaatg	tactaggtga	tcattttctg	cagatgacat	ggccgcacaa	1140
gtgagcgaaa	agggctacac	gtttgcgaat	ggtgctgagt	cttatgacgt	tgctgcagtt	1200
acgggggctg	tgaaacactac	agcaggtta	taacctgtgt	atgcgctatg	agacggttaa	1260
ttaaagcagca	gtgcagatga	taacgttact	cagacttact	acgagtgtgc	taattgtaac	1320
atttatgatg	acaaaaggctc	gtcactgtat	aaagctgcag	atggctctct	gaactctgaa	1380
gctaaaaggga	aattctgaag	aaccgcgcat	cccttgaaa	ctctggagca	agccatcagc	1440
tccatgcagca	aattccgctc	ctccctcggt	gcggttcaaa	acgctctgga	ttctcggtg	1500
accaactcga	acaaaccac	taccaacctg	tcgaagcgc	agtcctcgat	tcaggagacc	1560
gactatgcga	ccgaagtgtc	caatatgtcg	aaagcgca	tcattccagca	ggcggttaac	1620
tccgtgttgg	caaaaagctaa	ccaggatccg	cagcaggttc	tgctctctgt	gcagggttaa	1680

<210> 40  
 <211> 1146  
 <212> DNA  
 <213> Escherichia coli

<400> 40						
gcgctgtcga	cttctatcga	gcgcctctct	tctgggttgc	gcattaaacg	cgctaaagat	60
gacgctgtcgg	gcagggcgat	tgctaaacgc	ttcacttcta	acatcaaaag	tctgactcag	120
gcgcgacgta	acgccaacga	cggtatctct	ctggcgcgca	ccactgaagc	cgcaactgct	180
gaaatacaaca	acaacttgca	gcgtgttctg	gaactgacgc	ttcaggccac	tgactgactc	240
aactctgatt	ctgacctgtc	ttcaatccag	gacgaaatca	aatcccgctt	ggctgaaatc	300
gatcgtgtct	ctggtcagac	ccagttcaac	ggcgtgaacg	tgctggctaa	aaacggttct	360
ctgaattctc	aggttgggcg	gaatgatggg	cagaccatct	ctatcgattt	gcagaaaata	420
gactctcttc	cccttggttt	aagtgggttt	agtgtttgcg	gtggggcgct	aaaattcaag	480
gacacagtgca	cgaggttcgg	cgatgggtta	gccgcgccc	ttaaaagtga	tctggatgca	540
gtgcgcaacag	atatgtgtac	tgctttgggg	caaaaaggtta	atgcgaagtt	tttaagttg	600
cacaatatct	tagcaaaaga	tggtgcggca	actgagaact	atggtgttag	ctatggtagt	660
gataaattacg	ctgcatctgt	tgcatgatgc	gggactgtta	ctcttaataa	aacggatatt	720
acttattacgc	gcgggtgatat	taccggcgct	accaagaagt	atacgttgat	taaaagtgtc	780
gctaattctc	acggagaggc	cgttggtttc	gctaccgctc	agggtaagaa	tatgaaatt	840
acagatgggt	taaaaaacca	gtccactgct	gcaccaacgc	atatgtctca	gaccattgat	900



ctggatacgg	ctgatgaatt	tactggggct	tccactgctg	atccactggc	acttttagac	960
aaagctattg	caacaggtga	tactttccgc	tctccctcgc	gtgcggttca	aaacgctctg	1020
gattccgcag	tcaccaacct	gaaccaact	actaccaacc	tgctctgaag	gcagtcocgt	1080
attcaggagc	ccgactatgc	gaccgaagtg	tccaatatgt	cgaaagcgca	gatcatccag	1140
caggcc						1146

<210> 41

<211> 1506

<212> DNA

<213> Escherichia coli

<400> 41

atggcacaa	tcattaatat	caacagcctc	tcgctgtag	ctcaaaaata	tatcaacaag	60
aaccagctcg	cgctgtcgag	ttctatcgag	cgctgtctct	ctggcttgcg	tattaaacag	120
gcgaaggatg	acgcagcggg	tcaggcgatt	gctaaccgtt	ttacttctaa	tattaaaggg	180
ctgactcagg	ctgcacgtaa	cgccaatgac	ggattttctc	tgccgcagac	cactgaagcg	240
gcactgtctg	aaatcaacaa	caacttgcag	cgctgtcggt	aactgacccg	acaggcgaca	300
accggaacga	actcgaatc	tgacctgtcc	tctatccagg	acgtaataca	atccgtctgt	360
gaagagattg	acgcgtatc	cgccagactc	cagtttcaag	gggtgaatgt	gctggcaaaa	420
gcagcgacca	tgaaaaatca	ggtaggcgcg	aacgatggtc	agactatctc	tatcgattct	480
aaaaaaatcg	actcttcaac	cctgggctcg	acgggttttg	atggttcgac	gaaagcgagt	540
attttctacg	cagcagtaac	ggggggcgca	acgaccactt	atgctgtagt	cgccgttgca	600
attgatctcg	gaacggatat	tagcggatatt	gctgtctgat	ctgcgttagg	aacgatcaat	660
ttcgataata	caacaggcaa	gtactacgca	cagattacca	gtcgggccaa	tccgggctct	720
gatgggtgctt	atgaaatcca	tgtaaatgac	gggagtggtt	ccttcaactg	agcagcgagt	780
gataaaacag	cggtgtctgc	tccgggtact	gctctgacaa	gcggtaaaag	tcagactgca	840
accaccaacc	caggtacggc	tggtgatgtc	actgcggcta	aaactgctct	ggctgcagca	900
gggtgtcgaca	cgagtggcct	gaaactgggt	caactgtcca	acacgggatt	cgcaggtaaa	960
gtgaccaacg	tggtttacgg	cctgcagaat	gacagcggca	ctatctttgc	aaccgactac	1020
gatggaccaca	ctgtgaccac	gcggggcgca	gagactgtga	cttcaaaaag	tgcttcgagt	1080
aaacgacacca	ctgcggctgt	cacactgggt	ggctctgatg	gtcgtgatct	tcgcgatctc	1140
gcgcgtcgacg	gcataaacgta	cggtgcgact	gcaactgaatg	gtgctgcctc	ggatgatgca	1200
aataaacacg	ttaaatctgt	tgacagcaac	gctaaaocgt	tgcaaaaacg	ctctgattcc	1260
attgctgatgg	tcgacaaatt	ccgctctctc	ctcggctcgg	tgcaaaaacg	ctctgattcc	1320
gcagtcaccca	acctgaacaa	caccactacc	aaocctctcg	aaocgcagtc	ccgtattccg	1380
gacgcgcgact	atcgagaccga	agtgctccaa	atgtcgaaa	cgagattatt	ccagcaggca	1440
ggtaactctcg	tgctgtccaa	agctaaccag	gttcgcgagc	aggtctctgc	tctgctgcag	1500
ggttaa						1506

<210> 42

<211> 950

<212> DNA

<213> Escherichia coli

<400> 42

aaacaaaaac	agtcgcgct	gtcgactctt	atcgagcgcc	tctcttctcg	tctgcgtatt	60
aacagcgctta	aagatgacgc	cgcgggccag	ggatttgcta	accgctttac	ttctaaacct	120
aaaggctctga	ctcagcgccg	acgtaaacgc	aacgacgcta	ttctcttgcc	gcagcagcgt	180
gaagggcgcg	tgctcagagat	taacaaacaa	ttgcagcgta	ttcgtgaact	gaacctcag	240
gcctctctacg	gcacgaactc	tgattccgac	ctgtcttcta	ttcaggacga	aatcaaatcc	300
cgctcttgatg	aaattgaccg	tgatctcggt	cagaccocag	tcaacggctgt	gaacgtgctg	360
tgtataaacg	attcgatgaa	gattccagatt	gggtcccaatg	ataaacagac	gatcagcatt	420
ggcttgcaac	aaatcgacag	taccactttg	aatctgaaag	gatttaccgt	gtccggcatg	480
gcggagattca	gcgcggcgaa	actgacggct	gctgatggta	cagcaattgc	tgctgcggat	540
gtcaaaagtg	ctggggggtaa	acaagtcaat	ttactgtctt	acactgacac	cgcgctcaac	600
agttactgaa	atgcggctcg	tgattctgca	accggttaat	acatggaagc	cactgtatgc	660
attactcgctt	cgcgggcgcg	ggtaactggt	ggtagcggcg	aaagtgcggg	agcgcgtaca	720
ccgcatccgt	taaaagcact	ggatgcgcga	atcgcttaac	tcgacaaatt	ccgctctctc	780
ctcgggtcgcg	tcaaaaacg	cttgattctt	cgggctcaaca	acctgaacaa	caccaccacc	840
aaactgtctg	aagcgcagtc	ccgtattccg	gacgcgcgac	atgcgaccga	agtgctccaa	900
atgctcgaaag	cgcagattat	ccagcaggcc	ggtaactccg	tgctggcaca		950

<210> 43  
 <211> 1707  
 <212> DNA  
 <213> Escherichia coli

<400> 43  
 atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaataa tatcaacaag 60  
 aaccagctctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120  
 gcgaagagat acgcagcggg tcaggcgatt gctaaccggt ttacctctaa cattaaaggt 180  
 ctgactcagg ctgcacgtaa cgcacaacag ggtatttctg ttgcacagac cactgaagcg 240  
 gcgctgtccg aaatacaaaa caacttacag cgtatccgtg aactgacggg tcaggcttct 300  
 accggggacta acccgctatt ggtatcgag tccattcagg acgaaatcaa atccgctctg 360  
 gacgaaattg acgcggtatc cggctcaacg cagttcaacg gtgtgaacgt gatgatctg 420  
 gacggttcga tgaaaattca ggttggtgag aatgacggcg agactatcac gatgatctg 480  
 aagaagaattg actcagatag gctggggctg aatggtttca acgttaatgg caaaggcact 540  
 attgcaacaa aagctgctac agtcagcgat ctgaccgctg ctggtgcaac ggtgaacagg 600  
 ccttatgtctg tgacacacaa caatacagca ctacagcgta gcgatgcaac ggtcgcctg 660  
 aaaaaccggag atacagttac tactactggc tggagtgtg ctgatctata ttaatgatcg 720  
 gctaaaggga acttcaccac tcaagcaaca gttgcagatg gcgatgttgt taactttgag 780  
 aatactctga aaccagcggc tggcactact gcactcaggt tttatactcg tagtatctgt 840  
 gatgtgaagt ttgatgtaga tggtaattgg gatgtgacca tcggtggtaa agcccgctac 900  
 ctggagcgca ctggttaacct atctacaaac aaccccgcca ttgcatcttc agcgaattg 960  
 tccgatctgt ttgctagcgg tagtacctta gcgacaactg gcttatccca gctgtctggc 1020  
 acaacttata actttggtgc agcggcaact ctggcgtaa cctacaccaa aactgtaaag 1080  
 gctgatactg tctgagcac agtcagaggt gctgcaacgg ctaacacagc agttatctgt 1140  
 gcgacaataa agtataatca aggtatcag tctgcaacgg cgctctctcg tgggtgcaag 1200  
 actaatgggtg cgtgtaattc gaatgacacc tatactgagc cagacaaga gctcagact 1260  
 accgcatctt accatataca ctacaaacgtc gataaggata ccggtacagt aactctagct 1320  
 tcaaatggcg caggtgcaac tggtaaaatt cgagctactg ttggggcgac cctgtatgtt 1380  
 aactctacag gcaaacgtac cactgaaacc accagtgcag gcactgcaac caaagatcct 1440  
 ctggctgccc tggatgaagc tatcagctcc atcgacaaat tccgtgtcac cctgggtcgt 1500  
 atccagaacc cctgtgattc cgcggttacc aacctgaaca acaccatac caactgtccc 1560  
 gaagcgagat ccgtatttca ggacgcgac tatggcgac aagtggtcca catgtcgaaa 1620  
 gcgcagatta tccagcaggc cggtaactcc gtgctggcaa aagccaacca ggtaccgcag 1680  
 caggttctgt tctgctgca gggttaa 1707

<210> 44  
 <211> 1720  
 <212> DNA  
 <213> Escherichia coli

<400> 44  
 atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaataa tatcaacaag 60  
 aaccagctctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120  
 gcgaagagat acgcagcggg tcaggcgatt gctaaccggt ttacctctaa tattaacagc 180  
 ctgactcagg ctgcacgtaa cgcacaacag ggtatttctg ttgcacagac cactgaagcg 240  
 gcgctgtccg aaatacaaaa caacttacag cgtgtgggtg aactgacggg tcaggcgacc 300  
 accgggtacca actcccgatc tgatctggag tctatccag cagaaatcaa atccgctctg 360  
 gacgaaattg acccggtatc cggctcagat cagttcaacg gcgtgaacgt actggcaaaa 420  
 gacggttcca tgaaaattca ggttggtgag aatgagggcg agaccatcac gatgatctg 480  
 aagaagaattg actcttctac gttgaaactg actggtttta acgtgaattg tctggttct 540  
 gtggcgaaata ctgcggcgac taagacagaa ctggctgctg ctgctgcgcg agcggtgata 600  
 actctctgctg tccgtactga cggcggtgacc aaataacagc tagacgcagg gcttaacaaa 660  
 gccacagcag caaacctgtt tgcaaacctt gcagatgggt ctgttgttta tgcagcatt 720  
 tccaaacgggt ttggtgcagc agcagccaca gactacacct acaataaagc tacaatgat 780  
 ttcacttcca atgcagcat tgctgtggtg gctggcgccg ctgatagtaa cagcgacct 840  
 ctgcaatctt tctgactccc aaaagcaggt gatcacgata acctgacgt caaaatcggt 900  
 acgacatctg ttaatgttgt tctggcgagc gatggcaaaa ttaacgcgaa gatggctca 960  
 gctctgtgata tgcactcaac gggtaacctg actcagaaca gcgcaggcac tgtaacagca 1020  
 gcaaacctgt atgactgac caaaaacctat gatgcgacag gagctgtgg ttgtgatac 1080  
 acgacgcgag atggcgcaac tatctctctg gcaggctctg ctaacgcggc aacagctact 1140  
 caatcagggt caattacact gaanaatggt cgtatcctgt ctgatctct cagctctcgt 1200  
 gcgaaaggta ctgttatcaa tgttgataat ggtgctgat atatttctgt tagtaaacac 1260  
 ggggtgtcgt actcagcagg gtgcgcctac ttatactgat gctgatgata aatcaacgac 1320  
 aaccaaccac gttgatatt tcttgcaaac tgatggcagc gtaaccaatg gttctggtta 1380  
 aggggttctt accgatgcag ctggttaaat cactacagac gctgcacaaa aagccgcaac 1440  
 caccaccgat ccgctgaaag ccttctgatga cgcaatcagc cagatcgata agttccgttc 1500

atccctgggt	gctatccaga	accgtctgga	tccgcgggtt	accaaaccga	acaaaccacc	1560
taccaaacctg	tccgaagcgc	agtcgccgat	tcaggacgcc	gactatgcga	cogaagtgtc	1620
caatatgtgc	aaagcgcaga	tcctccagca	ggccgttaac	tccgtgtgtg	acaaagctaa	1680
ccaggtaacc	cagcaggttc	tgtctctgct	gcagggttaa			1720

<210> 45

<211> 14516

<212> DNA

<213> Escherichia coli

<400> 45

gactctgatgg	cogtagggcg	ctacgtgctt	tctgctgata	tctgggctga	gttgaaaaaa	60
actgctccag	gtgctctggg	acgtattcaa	ctgactgatg	ctatgcgaga	gttgagctaa	120
aaacagctgt	ttgatgccat	gctgatgacc	ggcgacagct	acgactgcgg	taagaaagtc	180
ggctatatgc	aggcattcgt	taagtatggg	ctgcgcaacc	ttaaagaagg	ggcagaatgc	240
cgtaaagaca	tcaagaagct	actgagtgag	tagagattta	cagctcttgg	ggacgataag	300
ccagaaaaaa	tagcggcagt	taacatccag	gcttctatgc	tttaagcaat	ggaggtttac	360
tgcgcttttt	tatgaaaaat	gaccaataat	acaaagttaa	ccctaccagt	ttaactgtct	420
ttttgttggg	tttttctctg	ttcttggtgc	catttggtaa	gacaattagc	gtgagtttta	480
gagagttttg	cgggactctg	cggaaactgc	cacattcttg	gcatttagtt	agtgctactg	540
tgcgtcttaa	tagcggggcg	gtagcttgcc	taattaattt	ttaacgtata	catttattct	600
tggcgtctat	agcaaaaaaa	gtcaatccga	ttaaaactct	tttccattag	gtaaaaaggt	660
gtttgtagtc	gtctcaggaa	attggttttg	gtagtagtac	ttttccaaat	atccattttt	720
cgatttagatg	ggcagttgat	gttactatgc	tgcatcacata	tcaatgtata	ttatttacct	780
ttagaatgtg	tatgaaaaaa	aatagtgctc	ataggccaag	tagcgtcaat	gatgttaagg	840
ttcagaaagg	caataatcat	gaatttagtg	aggcaagggt	ataatgtata	ctgtctagag	900
aatgattttt	ccactgaaga	tcttaaaagt	cttttgcgat	gggggtgtaa	gggggtttaa	960
ttctctctta	actcaaaagg	tattaactct	tttaaggata	taattgctgt	ttatgaacta	1020
aaaaaaattc	ttaagatatc	ttccccagat	attgtatttt	catattttgt	aaagccagta	1080
aattttggaa	ctatctcttc	aaagtgtcga	aaagtcgcaa	ggatgtgttg	aatgattgaa	1140
ggcttaggta	atgcctttac	ttattataag	ggaaaagcaa	ccacaaaaac	taaatgatga	1200
aagtggtatc	aaattctttt	ataatcagta	gcatcacgta	tgcttgatga	tttagttcta	1260
tttaactcatg	atgataaaaa	agattttaac	gaatcagata	ataaagaagc	taaggttaaca	1320
gtgttaggtg	ggatttgatt	ggactctaat	gagttttcat	aaaggaatga	accgaagaag	1380
aaaaatccct	ttatttttat	agcaagggtta	tcaagagaga	aaaggatatt	tgagtttatt	1440
gaagccgcga	agtcctgtta	gacacactat	ccaagtcttg	aaaattgtaa	tttagagagt	1500
tttgagagta	ataatccttt	ctcatatcaa	aaagtgtgaa	ttgaatcgct	aaagaaagaa	1560
catgacgtta	tttatcctgg	tcattgtggg	aatgttcaag	atcggttaga	gaaaagttct	1620
gtttttgttt	tacctacatc	atatcgagaa	ggcgtcccaa	gggtgatcca	agaagctatg	1680
gctatttgta	gacctgtaat	aacaaactaa	gtacctgggt	gtagggatat	ataaatagat	1740
ggggctcaatg	gcttttgatg	acttcatttt	gaatttaatt	tactggcgca	aaaaatgaaa	1800
gaaaaaaact	tgtagtcttt	tgaaaaaaat	gaattggggc	ttcgtggaa	gaagtttgca	1860
aatgattttt	gacttgagca	gaattatttt	ataattcaat	ctgaaaaata	aaggctttta	1920
ttatgaataa	agtggtcatta	attactggta	tcactggggc	agatggctcc	tattttggag	2040
aattatttgt	agaaaaaagt	tatgaagttc	atggtatttaa	accgcgttca	tttctattta	2100
atactgagcg	agtggtacac	atctatcagg	attcacattt	agctaactct	aaactttttt	2160
tacactatgc	cgatttgaca	gatacttcca	atctgaccgc	tattttaaaa	gaagttcaac	2220
cagagaaggt	tacaattttg	ggggcgatga	gccatgtagc	gggtatcatt	gagtcaccag	2280
aatacactgc	tgatgttgat	gcgataggaa	catgtgctct	tttgaaagct	atcaggatat	2340
tggggctgga	aaaaaagaca	aaatttttat	aggcttcaac	ttcagaagct	tatggttttg	2400
ttcagaagat	tccacaaaaa	gagactacgc	cattttatcc	acgttcgcct	atgtctgttg	2460
caaaattata	tgctctattg	atcactgtta	attatctgta	gtcttatggg	atgtttggct	2520
gcaatgtagt	tgccttttaac	cacgaatcac	ctgcgcgtgg	cgagacattt	gttaactctga	2580
aaatacattc	cgggatagca	aatattgtct	aaggtcttga	taaatgctta	tacttgggaa	2640
atatgcacgc	tgtctgtgat	tggggacatg	tcaaggatta	tgtaagattg	caatgtagta	2700
tgctgcagca	agaaactcca	gaagattttg	taattgtcat	aggaattcaa	tattctgtctc	2760
gtgagtttgt	caaatggcgc	gcagagcaag	taggcataga	gttagcattt	gaaggtggagg	2820
gagtaaatga	aaaaggtrgt	gtgttttcgg	tcaatggcgc	tgatgctaaa	gctgtaaaac	2880
cggcgagatg	aattatatact	gtagatccaa	ggtattttag	gcctgcagaa	gttgaaacct	2940
tgcttggcga	tcttactaat	gcgcataaaa	aattaggatg	gagccctgaa	atcacattgc	3000
gtgaaatgtg	aaaaagaatg	gtttccagcg	atttagcaat	acgcgaaaaa	aacgtgttgc	3060
tgaaagctaa	tacatttgcc	actaatattc	cgcaagaata	aaaaagataa	tacatttaaa	3120
aattaaaaaa	ggtctagatg	tatttagtat	cattattttt	ttttgggtga	ctaatgttta	3180
ttacatcaga	taaatattga	gaattatata	agttagtctc	attagtatca	attgactctg	3240
taattgaaaa	cgagaatggt	gaattattat	ttggtcttag	gaataatgca	cccgcaaaaa	3300
attatttttt	tgttccaggt	ggtaggattc	gcaaaaatga	atctattaaa	aatgctttta	3360

aaagaatc	atctatgaa	ttaggtaaa	agtagtgat	ttcaggaa	gtttttaa	3420
gtgtatg	acattctct	gaagtaggt	ttttttctg	aggcgagg	acattata	3480
tagtgctt	ttacacact	aaagtcctt	aaagtgatt	gaatctcca	gatgatca	3540
atcggaata	ttcttggtc	actaaacac	aaataaatg	taacaaga	gttcataac	3600
attcaaaaa	ttctttttg	taatttttat	taaaaaata	tatcgagag	aattgtatg	3660
ctcaatgtc	ttacccctg	attattgct	gaggaaacg	aaagcgtct	ttgccgtct	3720
ctcgagtat	ataccctaaa	caatttttaa	atttagtct	ggattctaca	attgttgcaa	3780
caaccaatt	cggtttgga	ggcatcgaa	gggaaatcg	aatgtgtatc	tgcaatgag	3840
atcaccgatt	tattgtagca	gagcaattac	gacagatcc	taagctaacc	aaatatatat	3900
tacttgagcc	taaaagccg	aatactgcac	ctgcctatgc	tttagctgtc	tattcgctc	3960
agaagaata	gcttcaatg	gaccttttat	tattagctac	ttgcggcagc	caactctata	4020
atgattgaaa	agcatctcg	gagtcataaa	taaaaagct	gcgcgtatgc	actttctgga	4080
agttgatga	attttgaaat	attccggaca	cggttaatac	tggttatgca	taattattgca	4140
gaagttcttc	agctgatact	aataaagaat	tccagcatac	taagtgtgcg	gaggtttgag	4200
aaaaaacaga	tggttaaaac	gcacaggaat	attattcgag	tgggaaattc	tactggatac	4260
cgcggaattgt	tttatttcgc	gccagtaaat	attctgatga	actacggaaa	tttagaacag	4320
atatttatca	tagctgtgaa	gtgcacccg	ctacagcaaa	tatagatatg	gaattttgtc	4380
gaatttaacga	ggctgagttt	attaattgtc	ctgaagagct	tatcgattat	gctgtgtatg	4440
aaaaaacaaa	agacgcctg	gtcttccga	tagatatgtg	ctggaaatgc	tttgggtctt	4500
gggtcatcat	ttgggatata	agccaaaaag	attgcactgg	taattgtctg	catggggatg	4560
tgctcaacta	tgtcggagaa	aatagtttta	tttactctga	gtcgaagtgc	gttgcgcagc	4620
tcggagtaga	taatttagta	attgtccaaa	ccaagagatg	tgtactgggt	gcggagcgtg	4680
ataaagtcca	aaagtgttaa	aacatagttg	acgatctaaa	aaagagaaaa	cggtgctgaa	4740
actacatgca	tgctgcagtt	tttcgcccc	gggttaaat	cgatgcata	gaccaaacgc	4800
atagatgata	agtaaaaaa	ataatagtta	accaggagag	aggggttagt	ttaaaggagg	4860
atcatcatag	ggcagagcat	tgattgtgtc	tatccggatc	ttctaaagt	taactaggaa	4920
gtagagttaa	actattatgt	tctaagtagt	ctatatatac	ccctcaggga	gcaaaaatga	4980
gtcttgagaa	tcacaggcga	atacctttgc	atctaattga	agtaagttct	gggtattaac	5040
tgtaactcga	tgatatagtg	cgttttactg	acagatataa	catgataaaa	ttctctaaag	5100
gagattgata	aatatgaata	aaataaactg	ctcaaaagca	tatgatatac	gtgggcctct	5160
tggtgtctgta	ttgaatgatg	aaatagcata	tatgaatgtt	cgcgcttatg	gtgagttatg	5220
taaaactcca	actgtagttg	tgggaggaga	tgctcgctta	acaagtctga	gtttaaagaa	5280
actactctca	aatgggctat	gtgatgcagg	cgtaaatgtc	ttagatctgt	gaatgtgtgg	5340
tactggaagag	atcatatttt	ccacttggtt	tttaggaatt	gatgggtgga	tcgaggttaa	5400
tgcaagccat	aatccaattg	attataatgg	aatgaattga	gttaacaaag	gttaactgac	5460
aatcagcagt	gacacaggtc	tcaaaagatt	acaaacaatt	gtagagagta	ataattcttg	5520
agagctcaac	ctagaaaaaa	aaggggaatt	acaaataat	tccaccocag	atgcctacat	5580
aaatcatttg	atgggctatg	ctaactcgca	aaaaataaaa	aaaatacaaa	tagttgtgaa	5640
ttctgggaaat	gggtgcagctg	gtcctgttat	tgatgctatt	gaggaatgct	tttttcgaaa	5700
caactctcgc	attcagtttg	taaaaaata	taatacacc	gatgttaatt	ttccactatg	5760
tatccctgat	ccattactac	ctgagtgca	agaagatacc	agcagtcg	ttataagaca	5820
tagtgcgtat	tttggtatg	catttgatg	tgattttttt	tgattttttt	tttttgata	5880
aaaatggcaaa	tttatgtga	gatatacat	tggtggttta	ttagcggaag	tttttttagg	5940
gaaatatccca	aacgcaaaaa	tcattcatga	tcctgcctct	atatggaaata	cttttgatat	6000
cgtagaaagt	catggttgga	tacctataat	cgctaaaaac	gggtcatgct	acattatgca	6060
aagaattcgt	gaagaggatg	cgttatatgg	cggcgaaatg	agtgogcatc	attatttttaa	6120
agatttttga	tactgcgata	gtggaatgat	tccttggaat	ttaattttgt	aaacttttga	6180
tctgacaaat	aaaaaattag	gtgaactgg	ttgtgtgtgt	ataaacgact	ggccggcaga	6240
tgaggaaata	aactgtacac	tagacaatcc	gcaaaatgga	atagatatat	tatttttaact	6300
ttcaaaaagt	agtccttgta	ctgttgatta	caactgatga	ttaaactatg	ttttctctga	6360
ttggcgtttt	aatgttagat	gtcacaatac	agaacctgta	gtacagattg	atrgagaatc	6420
taggaataat	gctattttct	tcgcagaaaa	aacagaagaa	attctgaatt	ttatatcaaa	6480
ataaatttgc	actgagtttc	ataatgggaa	caagaaatag	atgaaagtac	ttcrgactgt	6540
ctcaactggc	acgtttggtg	agaatatatt	agagcatgat	agtgcaagta	aatataaatt	6600
acttactcca	accagctctg	atttgaattt	attagataaa	aatgaaatgt	aaaaatttcat	6660
gctttatcaac	attccgagct	gtattatata	tcgcgcggga	ttagtgtgag	cgattctcgt	6720
aaatataaag	agccgctttg	attttctgga	aaaaaattgt	cgatgtgggt	taaaatttagt	6780
ttccgtctgca	aaaaaactag	gatatcaaga	agtgcttaac	ttgggtagtt	ctatcgatgta	6840
ccccaaaaac	tttgaagagg	ctattctctg	gaaagctctg	ttaaactggg	agcrgagaag	6900
aactaatgag	ggatatgcta	ttgcgaaaat	tgctgatgca	aaagcatcgt	aatatatact	6960
aagaaaaaatt	tctaatattt	tttataaaac	aattatccca	tgtaatttat	atgggaaaata	7020
tgataaaaa	tgatataact	cttcacatat	gattccggca	gttataaaaa	aaatccatca	7080
tcgcaaaaatt	ataaattgtc	caagagatga	aatttggggg	gatggttaatt	cgccgcgtga	7140
gtttattgat	gcagaagatt	tagctgatct	tatttttttt	gttattcccta	aaaragaatt	7200
catgcctaat	agttgaaatg	gtgttttagg	ttacgattat	tcaatttaag	actattataa	7260
gataaattgca	gaagaaattg	cttatactgg	gagttttttc	catgatttaa	caaaaccaac	7320
aggaatgaaa	cggaagctag	tagatatattc	attgcttaatt	aaaatttggt	ggccaagtca	7380

ctttgaactc	agagatggca	tcagaaagac	ctataattat	tacttggaga	atcaaaaataa	7440
atgattacat	accactctgc	tagtaatact	tgggatgaat	atgagtatgc	agcaatacacg	7500
tcagtaaatg	ctacaaaaat	gtttaccact	ggtaaaaaag	ttgagttata	tgagaaaaaat	7560
tttgcgtagt	tggtttggtag	caaatatgcc	ataagggtca	gctctgggtc	tacagctaatt	7620
ctgttaacga	ctgtgtccct	ttcttccact	acataaccac	aaccttaaaag	agggtgatga	7680
ataatagtac	ctgcagtgctc	atgggtctacg	ctctgcacaa	ttgatatacga	tattttgaaa	7740
aagggtgaagt	ttgtcgatat	caataaagaa	actttaaaata	tttatagtag	taactctaatt	7800
aatgtctatt	cagataaaaac	aaaagacaata	ttgacagtaa	attatttagg	agaagataact	7860
gattttgcgaa	aaaataaagt	gataataaat	aataagggata	ttactcttact	agatgaagg	7920
ttgtgagtcca	tggggcgcggt	ctttccaaaat	aagcaggcag	gcacattccg	agtattgggt	7980
accttttagtt	cttttttactc	tcacatata	gctacaaagg	ggagggggctg	cgtagtgtact	8040
gatgatggaag	agctgtatcca	tgatatgtgtg	tgctctccgg	ctcatgggtg	gacaagaatat	8100
ttaccaaaaag	agaatatgggt	tacaggccact	aagagtgatg	atattttcga	agagtcggtt	8160
aagtttgtttt	taccagggata	caatgtttcgc	ccacttgaaa	ttgttgggata	tattgggata	8220
gagcaactta	aaaaggttacc	aggtttttata	tcacccagac	gttccaatgc	acaatttttt	8280
gtagataaaat	ttaaagatcca	tcacattccct	gatatacaaaa	aagaagtttg	tgaaaagtgc	8340
ttgtttgggtt	tttcccttcgt	tataaaaggag	ggagctgccta	ttgagaggaa	gagtttagta	8400
aataactctga	ttctcagcagg	cattgaaatgc	cgaccaattg	ttactctcaa	ttttctcaaa	8460
aatgaacgctg	tttttagtta	ttttgtctac	ttctgacatg	atacggtagc	aaatgcgcga	8520
tatatagata	agaaatgggtt	ttttgtcggga	aaccaccaga	tcactttgt	taatgaanaa	8580
gattactctac	aaaaagattt	aaaaataacta	acgaggcact	ctatttcgaa	tagagtgcct	8640
ttaaagatggg	attaacagctg	aaaaaaaattt	tagecgttgg	ctattctaaa	gtactaccac	8700
cggtttatgga	acaggtttgtc	aatccaattt	gcactttcat	ctaacaccac	ctaatatcca	8760
accacctcggg	taagcaaaagc	tatggtaatt	ggattttatt	aattactatt	tactcttttt	8820
ctcagttaaat	ctgtggaggga	tggttcgcgat	ggaattgcaa	aatcatctga	gaacagagaa	8880
ttctttagtta	tttttcaaaa	aaaaatgctt	tacgtccaat	ttctcataat	ttttccattg	8940
ttattatcgc	tttttcgggta	ttgattttctt	ttctctattt	aagratattgt	ttcttcgagt	9000
ttgcggaggaa	taattctctca	ttcttattcgc	cgattattatt	tttgggtttt	tttaaggaaa	9060
ttgataaatt	ttatttagtgt	ggcgtaaaaag	cttatgtaaaa	attcaatgta	tcagtgtttt	9120
ttgaagttaat	tacaagagtg	ctctgggctt	ctatagtaatt	atatggcatt	tacgaaaagt	9180
caactcttata	ttttacatgt	tttagccttta	ccattaaaagg	tatgtcaaaa	tatatctctt	9240
tattgtctgaa	tattaccgggt	ttgttcatcat	atctcaattt	taataagagtt	ggagttctga	9300
atttgtttaaa	ttgagtcaaaa	ttgatgttttc	ttcaattaac	ttgttggcgt	ttcaactagtt	9360
ttgtttgctga	ctcgtgtaata	ccattgttttc	tatctgtccag	taactcggct	ctctatgcgc	9420
cttgcgttcca	actagctcaaa	ttgatgttttc	ctctctctgc	gtctgcataat	caaatattac	9480
tacocatgtt	tgctagaaatg	aaagcatctc	acacatttcc	ctctaattgt	ttttttaaaa	9540
ttctgcttgt	atcactaatt	tcgtttttgc	ttctgtcttc	gttatctttt	tttggctcgt	9600
atatattatc	aaataggata	aacctacact	ttgcacactga	aaatataaa	ttatgcaaaa	9660
tttttagctat	aagtacagtt	ttattgtcaa	tgatgacatc	ttttcatttc	ttgttattag	9720
gaatttggtaa	attcaagctt	gttgcaaaat	taaatctggg	tgacaggctc	gcacttgctg	9780
cttcaacgtt	aatcgcagct	cattatggcc	tttatgcgat	atctatggta	aaaaataat	9840
atccggcttt	tcgaattttat	taccttttatg	tagcttttgt	ctattttaat	agagcgaaaa	9900
atgtctattg	attttactttt	ttcaattact	gaaatgcgaa	ttgtttttt	tttccattat	9960
tacataattta	ctcaactggt	gttaatgcgg	aggatctatt	tagataaaaag	tttttaatt	10020
ctttttatgct	tgctctctttt	tttagttaatc	attcaacttc	ctgagcttaa	tgtaaacagg	10080
ttgttgcgatt	cttttaagtt	atcactgcct	tattgatgg	ttttatcgc	ttttcaaaaa	10140
ocgaaatatt	gcttgtgggtg	tattattgca	ttgttttttt	tgaaactcgc	atttaatttt	10200
ttatatttaa	agacattcga	taagttttagc	ttatttccct	ttaacttttt	ttatttgctg	10260
ttttacttgt	ttagattggg	aatttggtaat	ttaccgggtt	ataaaaaata	aaaaatttacc	10320
gggttgattt	ttctctttat	ataatagac	ataatgcagt	cattgttaat	aaatatagg	10380
gggcagattt	tatatctcgg	aatttgcact	ctgataactg	tggtttaaagt	taattttaaga	10440
aaaaagattc	catacttttt	tttaatgctg	ccagttttat	atgtaattat	tattggcttat	10500
atrggttttta	attatttcaa	taaaaggcgt	actttttttg	aacctacagc	aagtaataat	10560
gaacgtacgg	ggatgatata	ttatttggtt	tcacagctgt	gtattatat	atttccatgg	10620
atrggggaatc	taaaattctt	aaataacggc	ggacaataata	agacgttata	tggaattcca	10680
tcattataat	cttaatgacc	tcagtatttt	tttatacgg	ttttataagg	tattgttggtg	10740
ataaggagcat	gggttatcca	ttctatattt	tttgtttttt	ttaggagaat	atctttcttca	10800
ttattagaga	taaatgctcc	tttcattggt	gtaagtgtgt	ttgtactgtt	acaaagtgtg	10860
ttaatattata	catnaaaccc	ttttgtatgct	tttaatcgat	tgatttgcgg	gcttaccagt	10920
ggagttgttt	attgattttg	aaaaattaga	taagtatacc	tgtaattgaa	atttagagc	10980
ttccacttgtt	ttcaataatca	ttgcaactta	taattctgaa	ctgtgatag	ctaaagtgtt	11040
gcaatcggta	actaaatcaat	cttataagaa	tattgaaatc	ataataatg	atggagagatc	11100
ttcttgataaa	acgtctgata	ttgcaaaatc	gtttaaagac	gaccgaataa	aaatagtttt	11160
agagaanaag	ctgtgaaatt	atgtagcctg	gaataaagca	gttgatttat	ctattgggtga	11220
ttgggtgagga	tttatgtgtt	cagatgatgt	ttactatcat	acagatgcga	ttgcttcaat	11280
gatgaagggg	gttatgggat	ctaatggcgc	ccctgtgggt	tattgggagga	cagcgacga	11340
agggtcccgat	aggaacatat	ctggatttttc	aggcagtgaa	ttgttacaacc	tacaaagatt	11400

taagttttaat tattacaaat gtaattttacc attgcccatt atgagcgcaa tatattctcg 11460  
 tgattttcttc agaaacgagc gtttttgat ataatataaa attgttctgt accgtgatgt 11520  
 gttttctgaga tggttctatca ataggagtaa agagaagatca cctattttta ttaagtgcac 11580  
 gacccttatt gtagaatgg gatatggtgg ggtttcgact gaattttctt tcaagattaa 11640  
 aacrcgccta gaaagtttca ttgtacgcga aaagaataat atactctgtt taaacatca 11700  
 gctgtctctt agatattgcta aaattctcgt gatggtgagc atcaaaaaa tttttgcgaa 11760  
 taatgtttat aaattgaatg ataacgggta tcaattcccta aagaaaaatca agaatataat 11820  
 atgaagattg tttatataat aacccgggctt actgtgtgtg gagcogaaca ccttatgacg 11880  
 cagrtgacag acaaaatggt tatacgcggg catgatgtta atattatttg tctaactcgt 11940  
 atactctgagg taaagcgaac acaaaatatt aatattcatt atgttaatat ggataaaaaat 12000  
 tttagaagct tttttagagc tttattttcaa gtaaaaaaaa taattgtcgc cttaaagcca 12060  
 gatataatca atagtcatat gtttctatgt aatattttta tgcgttttat taggtgctgt 12120  
 attccagcgg tgcctctgat atgtaccgca caacacaaaa atgaagggtg caatgcgaag 12180  
 atgtttttgt atcgactgag agtaaaagggt acocctaaaa ataaaaatgt agagattccg 12240  
 gctgttccaag agtttatagc gagaagaagg acocctaaaa ataaaaatgt agagattccg 12300  
 aatttttatta atacaaataa atttgatttt gatatttaag tcagaaagaa acgcgcagat 12360  
 gcttttaatt tgaaaagcag tacagcagta atgtctgcag taggaagact tgttgaagca 12420  
 aaagactatc cgaacttatt aaatgcataa aatcaatttg tcttttcaa aatattcaat 12480  
 tgtaatgatt ttaattttgt tatgtctgac gatggtcgt taagaaataa atatttgat 12540  
 ttggtttgtc aatggaatct tgtggaataa gtttctctt tggggcaaaq aagtgaatt 12600  
 aaagaattaa tgtgtctgc agacttttt gttttgagt ctgagtgga aggttttgt 12660  
 ctogttgtg cagaagactt ggcgtgtgaa gctccogtt ttgtatgga tctgtgtgga 12720  
 gttaaagaag tgcgttgacc ctataatgat gttatccctg tcaagtaatca tattctgtt 12780  
 cgagagaaaa tgcgtgagac acttaaaaaa gatgataacg caagaaaaat aataggtatg 12840  
 aaaaaatagag aataatatgt tccaatttt tcaataaaaa cgaatgtgag tgaagtggag 12900  
 gccttatatt ttaaatattc caagcgtaat aatatattgt aatgaataa taagtttatg 12960  
 ctctgtagtc aaatgtttct ctatgctgtt tttttacttg ctccgtattt ttactttatg 13020  
 ctggattttg ttatatatca gtatacaact gctcaactt catctaatg acatccaag 13080  
 cgcgcgatcg tgcgcgggtg actacacctg accaggagat gtaattgcc aqnaacagat 13140  
 cgcgcgtcgc ttatagccag tgaatggggc caactctggc ctcaacatcg aaacccggcg 13200  
 ttatacgcgc ggcattctca accgctcccg cgagaagaat gaagaatttg ttgcgagaa 13260  
 ccccgataag aaacttgctc ctattatcac ggtgaaagag tctgcagat ctctgaaac 13320  
 cccagctcgt acctcgttaa tggtaaaagc agggcggtga actgatcgt ctatcgattc 13380  
 cctgaagctc tcttggtata aaggcgacat cattatgat ggtggcaaca cctctctca 13440  
 ggaacatcgc ctgtgtaaac tcttaactgtc cgcggaaagg ttaactctc tgcgtaccgg 13500  
 cgtgtccggc ggtgaagagg gcgcctctga aggccactct atcattgccg gtggccagaa 13560  
 agaagcgtat gagctgtgtg cgcctatcct gaccagaatt gctgcggttg ctgaagatgt 13620  
 cgaaccatgc ataactaca tccgtgctga cgtgcgggt cactacgtga agatgggtga 13680  
 caacgggtatc gaatatggcg atatgcagct gattgtctga gcctattctc tgcgtaaag 13740  
 cgcgtttaat ctgcttaacg aagagctggc acccaacttt accgagtgaa atgaaggaga 13800  
 gctaaatgcg taccctgatt acatcaccaa agacattctt acccaaaaa atgaaggagg 13860  
 cagccagcgc gttgatgtga tccgtgagca agctgcacga aaagccaccg taaatggac 13920  
 cagccagcgc tctctggtat tgggtgaacc gctgtcgtct atcaacogaat ccgtattcgc 13980  
 tgcgtacatc tcttctctga aagaccagcg cattgcgga tctaaagtgc tgtctggtcc 14040  
 cgaggctaaa ctgcgtggtg ataaagcaga gttcgttga aaagtccgtc gcgcgctgta 14100  
 cctggtgtaaa atcgtctctt atgcccaagg ctctctctca ctgcgtgccg tctgtgacga 14160  
 atacaactcg gatctgaact accggagaat cgcgaagatc ttccgcggcg gctgcatcat 14220  
 tccgtgcgag ttccgcaga aaattactga ccgttatgct gaaaacaaa catctgtctaa 14280  
 cctgtgtctg gtcctgtact taaaaatat cgtgatgat tatcagcaag cgtcgcgtga 14340  
 tgtagtggct gatcgtgac gaacacggtat tccggtacgc cctcgtctc cagcgtgac 14400  
 ctactacgac agctaccgtt ctgcggtact gccgccta atgtattcag cacagcgtga 14460  
 ttactctggt gcgcacacgt ataaacgcac tgataaagaa ggtgtgttcc acaccg 14516

<210> 46

<211> 1380

<212> DNA

<213> Escherichia coli

<400> 46

aacaaatctc agtcttctct tagctctgct attgagcgtc tgtcttctgg tctgcgtatt 60  
 aacagcgcaa aagcagatgc agcaggctcag gcgattgcta accgttttca ggcaaatatt 120  
 aaaggtgcta ccagggttcc ccgtaaacgcg aatgatggta ttctgttgc cgaagacct 180  
 gaaggtcgcg tgaatgaaat taacaaacac ctgcagcgta ttctgtaact ttctgttcag 240  
 gcaactgacg gtaactaact gcacagcgat ctttctctca tccagcgtga aattactcaa 300  
 gctctggaag aaattgaccg tgtatctcag caaaactcagt ttaacggcgt gaagctctt 360  
 gctgaaaaata atgaaatgaa aattcaggtt ggtgcctaag atggtgaaac catcactact 420  
 aatctgcgca aaattgatgc gaaaactctc ggctcgagcg gttttaatat cgaatggcgcg 480

cagaaagcaa	ccggcagtg	cctgatttct	aaattttaag	cgacaggtag	tgataattat	540
caaatattacg	gtactgtata	ctatactgtt	aattgtagata	gtggagtagt	acaggataaa	600
gatggccaacg	aagtttatgt	gagtgctgcg	gatgggttcc	ttacgaccag	cagtgtatact	660
caatttcaaga	ttgatgcac	taagctttga	gtggctgtca	aagatttagc	tcaaggtaat	720
aagattgtctt	cagaagggtat	cgaaattaga	ataccggcca	ctggcgctat	acctgccaca	780
ggtaaatcgtg	aatttaaccgc	caattgttgat	ggtaaggctg	ttgaattcac	tatttcgggg	840
agtgcctgata	ctactaggtag	tagtgcaacc	gttgccctca	cgacagccct	atacaaaaat	900
agtgcagggcg	aattgcctgc	aacaaaagtt	gaaaataaag	cagcgacact	actcgtattt	960
gattctgaacg	ctgccaaaga	aacaggaagc	acgttagttg	ttacagggtg	aacttactgt	1020
gttagctgcag	atgggtaaaac	gataacggag	actgctcttg	gttaacaata	agctatgtat	1080
ctagacaaat	cagaagggtgg	tagcccgatt	ctgtgtaaacg	aagatgcagc	aaaactcgtt	1140
caatctaccac	ccaaaccgctt	cgaaactatc	gacaaagcat	tggtctaaat	tgacaattct	1200
cgcttcgacg	ctgggtgcagt	acaaaaccgt	ttcgactctg	ccataccaca	ccttgagcac	1260
acgttaacaa	acctgccttc	tgcccgttag	gctatcgtag	actgcagcaa	ctgcagcgaa	1320
gtgtcttaaca	tgctctgtgc	gcagatccgt	caacaaggcg	gtacctctgt	tctggcacag	1380

<210> 47  
 <211> 1497  
 <212> DNA  
 <213> Escherichia coli

<400> 47	atggcacaag	tcattaatac	caacagcctc	tcgctgatca	ctcaaaaata	tatcaacaag	60
	aaccagctctg	cgctgtcgag	ttctatcgag	cgctgtctct	ctggctgtgag	tatttaacagc	120
	gcgaaggatg	acgcagcggg	tcaggcgatt	gctaaccggt	ttacctctaa	catttaaaggc	180
	ctgactcagg	cggcccgtaa	cgccaaacgac	ggtattctcg	ttgcgcagac	caccgaaggc	240
	cgctgtctcg	aaatcaacaa	caacttacag	cggtgtcggt	aactgacggt	acaggccact	300
	accgggtacta	actctgagtc	tgatctgtct	ttcttccagg	acgaataata	atcccgctctg	360
	gatgaaattg	acccgctatc	tggtcagacc	cggttcaacg	cggtgaacgt	gctggcgaaa	420
	aatggctcca	tgaaaatcca	ggttgccgca	aatgataaac	agactactac	tatcgatctg	480
	aagcagatgt	atgtctaaaac	ttctggccct	gatgggttta	cggttaaaaa	taacgatata	540
	gttaacacta	gtgcttccagt	aaactgcttt	ggtgtataca	ccacaacaaa	tattaaactt	600
	actggaattat	ccctttctac	ggaagcagcc	actgatactg	cggaataata	ccagcttcca	660
	attgaagggtg	tttatactga	taatggttaa	gattactatg	cgaaaataac	cggttggtat	720
	aacgtagggg	agtattacgc	agtaacctgt	gctaagtgat	gtacagtagc	aattggcgact	780
	ggagcaacgg	caaatgcac	tgtaactgat	gcaataacta	ctaaagctac	aactattact	840
	tcaggcggtta	caactgttca	gattgataat	actgcaggtt	cgcaaacctg	caaccttggg	900
	gctgttagct	tagtataaact	gcaggattcc	aagggttaat	ataccgatag	atacgcgctt	960
	aaagatacaca	atggcacaact	ttacgctgag	gatgtgaaat	aaactactgt	tgcgtttctt	1020
	gtctaaaacta	ttactataac	tgactcttcc	ggtgcccaga	gtttctcaac	cgcgcttcaa	1080
	ctgggcccag	atgctggcaa	acagcaagtg	ctcgatatgt	atggtaaaaa	atacgattct	1140
	gcggattttaa	atggccgtaa	ctgcacaaac	ggtttgactg	ctgggtggta	ggctctgact	1200
	gtctgttgcaa	atggttaaac	acaggatccg	ctgaaacgcg	tggaacgatg	tatcgatctt	1260
	gtcagcaaat	tcogtctctc	ctccggttgc	gtgcaaaacc	gtctggattc	cgcggtttac	1320
	aaactgaaca	acacactac	caactctgtc	gaagcgcagt	ccgtatttca	ggacgcgcag	1380
	tatgcgacgg	aagtgtccaa	tatgtcgaaa	gcgcagatca	tcacagcagg	cggttaactcc	1440
	gtgttgccaa	aagctaacca	ggtaccgcag	caggttctgt	ctctgtcgca	gggttaa	1497

<210> 48  
 <211> 1695  
 <212> DNA  
 <213> Escherichia coli

<400> 48	atggcacaag	tcattaatac	caacagcctc	tcgctgatca	ctcaaaaata	tatcaacaag	60
	aaccagctctg	cgctgtcgag	ttctatcgag	cgctgtctct	ctggctgtgag	tatttaacagc	120
	gcgaaggatg	acgcagcggg	tcaggcgatt	gctaaccggt	ttacctctaa	catttaaaggc	180
	ctgactcagg	ctgcagcgtg	cgccaaacgac	ggtattctcg	ttgcgcagac	caccgaaggc	240
	cgctgtctcg	aaatcaacaa	caacttacag	cgatctcggt	aactgacggt	acaggcttct	300
	accgggtacta	actctgagtc	ggatctggag	ttcattcagg	acgaataata	atcccgctctg	360
	gacgcaaatgt	acccgctatc	cggtcaaaac	cagtttcaacg	gtgtgaacgt	gctggcgaaa	420
	gacggtctgca	tgaaaatcca	ggttggtgag	aatgacggcc	agactactac	tattgatctg	480
	aagaaaatgt	actctgatac	gctggggctg	aatgggttta	acgttaacgg	aaaggttact	540
	atttcgcaaca	aagcggcaac	cattagtgtat	ctggcgccga	cgggggcgaa	tggtactaca	600
	tcaagcaata	gtgtgtccac	gacaaaagtc	aatgcctgtg	atgcagcagc	ctcagtttagc	660
	aaactcaaa	atggtagtgc	tggttcgggt	gctgctcaga	aataacttta	taacgcatcg	720

accaatgatt ttacgacaga aaatacagta gcgacaggca ctgcaacgac agatcttggc 780  
 gctactctgc aggcctgctg tggcgagagt caatcaggta catatacctt tgcaaatggg 840  
 aaagttaact ttgatgttga tgcaagcggt aatatcacta ttggcgcgga aaaggctctt 900  
 ttggcttggtg gagcgctgac tactaacgat ccacccggct ccactccagc acagatgtct 960  
 tccctctgtta aggcgcggga tgacaaaagt gccgcgtcaat cctcgattga ttttggcggg 1020  
 aaaaaatacgc aatttctgtg tggcaattct actaatggtg gcggcggtta attcaaaagc 1080  
 accggtgtctt ctgacgcgct tttggctcag gttaaaagcg atagtactgc taataatgta 1140  
 aaaaatcactt taacaatgg tctctgtgca ttcactgc atctccaaaa tgggtgtact 1200  
 ggctccgggg catcgaaatg agcttacatt gatagcgaag gcgaactgac aactactgaa 1260  
 tccctacaaca caaattattc cgtagacaaa gacacggggg ctgtaagtgt tacagggggg 1320  
 agcggtacgg gcaataacgc cgcaaacggt ggtgctcagg cttatgtagg tgcagatagg 1380  
 aatatcaacca cgaatactac tagtaccggc tctgcaacca aagatccact aaatgcgctg 1440  
 gatggagcaa tgcattccat cgcaaaattc cgtctctccc tgggggctat ccagaacctg 1500  
 ctggatccg cagtcacca cctgaacca accactacca acctgtctga agcgcagctc 1560  
 cagcagggcg cagccgacta tggcaccgaa gtgtccaa ca tgcgaaagc gcagatcatc 1620  
 cgtattcagg cgtactccgt gtgggcaaaa gctaaccagg taccgcagca ggttctgtct 1680  
 ctgctcgagg gttaaa 1695

<210> 49  
 <211> 1164  
 <212> DNA  
 <213> Escherichia coli

<400> 49  
 aacaaagaacc agtctgcgct gtcgagttct atcgagcgct tgccttctgg cttgcgtatt 60  
 aacagcagca aggatgacgc cgcgggtcag gcgattgcta accgttttacc ttctaacttt 120  
 aaagggcctga ctcagcgctgc acgtaacgccc aacgacggta ttctgtgtgc gcagaccacc 180  
 gaagggcgcg tctcgcaaat taacaaacaa ttacagcggt tgcgtgagct gactgtctag 240  
 cgccaccacgc gtacttaact tgaagctgac ctgtcttcta tccaggacga aatcaaatct 300  
 cgcctgggaag agattgatgc tgtttcaagt cagactcaat ttaacggcgt gaatgttttg 360  
 gctaagaatg tgaataatgaa catccaggtt ggggcaagt atggacagac tatcactatt 420  
 gatctgaaaa agatcgatcc atctacacta aacctctcca gttttgatgc tacaactttg 480  
 ggcaccagtg ttaaagatgg ggcaccatc aataagcaga tggcagtaga tctagttaa 540  
 tttaaagata aagcttcagg atcgttagtg acccttaaat tagttgagaa agacggttaa 600  
 tactatgtata atgacactaa aagtagtaag tactacagta cgaagtaga tactagtaa 660  
 ggtgaaatga acttcaactc tacaataatg aagtggaact ctctactgc agcgacgaa 720  
 gtaactactg ttggccgcga tgaataattg gatgctcttg cacttaaaac caaccactg 780  
 ctgtgctgtg ataaagataa aagcggcaat gatgcttata tcaatcgagc caaagatgta 840  
 acaactaacc aatcaacttt caatgccgct aatatcagtg atgctgtgtt ttattctatt 900  
 gggcctatca caacccgcgc aagcaattta acagctgacc cgcttaaggc tcttgatgat 960  
 gcaattgtcat cgtgtgataa atccgctctc tctctcgctt ctcttcagaa cgcctggat 1020  
 ctctgcattg ccaacctgaa caacaccact accaactcgt ctgaagcgca gtcccgatt 1080  
 caggacgctg actatgcgac cgaagtgtcc aacatgtcga aagcgcagat tatccagcag 1140  
 gccggtaact ccgtgctggc aaaa 1164

<210> 50  
 <211> 1818  
 <212> DNA  
 <213> Escherichia coli

<400> 50  
 atggcacaaag ctattaatac caacagcctc tcgctgatca ctcaaaaataa tatcaacaag 60  
 aaccagctcg cgcgtctgag ttctatcgag cgtctgtgct ttgctgtcgc tattcaacagc 120  
 gcgaaggatg acgcagcggg tcaggcgatt gctaaccggt tcaaccttaa catcaaaggc 180  
 ctgactcagg ctgacgtata cgctaacgat ggtatctctc tggcgacagc caagtaaggc 240  
 gcactgtctg agattaacaa caacttacaa cgtgtgctg agttgactgt acagggcacc 300  
 accggtacta actctgattc tgacctggct tctattcagg acgaataatca atccggttg 360  
 tctgaattgt acccgctatc cgggcagacc cagttcaacg gcgtgaacgt attgtctaaa 420  
 gatggctccc tgaataatca ggttggcgca aatgatggct agactatctc tatgcacctg 480  
 aagaaaaatt actctgatac tctgggtttg aatggtttca acgttaattg ttctgtatcc 540  
 attgcacaaca aagcgcacca aatcagtgac ttgactcgtc agaaagccgt agattcaacc 600  
 aatggtactt ataaagttac aactagcaac gctgcaacta ctgcatctca ggcattaaat 660  
 aagctgtctg atggcgatac tgtagatatt gcaacctatg ctggtgtgtac aagttcaacc 720  
 gttagttata aatacagcgc agatgcaggt aacttccagt ataaacaatc tgcaacaaa 780  
 acaagtctcg cggctggaac tctggcagat actctctccc cggcagctg ccagactaaa 840  
 accggtactt acaaggctgc tactggtgat gttactctta atgttgacgc aactggtaat 900



ctgacaaattg	gcggacagca	agcctacctg	actactgatg	gtaacctttac	aacaaacac	960
tccgggtggg	cggctactgc	aactctctaaa	gaagctgttta	ctcttgctgg	cgatggtaaa	1020
tctctggggga	acgcgggtac	tgctaccggt	actctggata	atactacgta	taatttcaaa	1080
cgctgtcgga	cgcttactga	tggtgtcggt	gtcatcgctg	ctgctgggtg	aacttataca	1140
gccactgttt	ctaaagtatt	catctctggca	caactcgcaat	ctgcgaagtca	ggcagcagca	1200
accgctaccg	acgggtgatac	tgctcgcaagc	atcaactata	aatcttggtgt	catgatcggg	1260
tccgctaccct	ttaaccaatgg	ttaaaggtact	gccgatggta	tgacttctgg	tacaactctac	1320
gtgctagacta	caggctgctaa	agctgtatata	gttgatggca	acaatgaaat	gaactccaat	1380
gcactctacg	tacagactta	ctctgtcaac	gcagataacg	ggcgagtaaa	agtggtatca	1440
ggatctcgta	ctggtaaaat	tgaagctggt	cgctggctgg	atgtcttatgt	aagcaaatgat	1500
ggcaaaattaa	cgacagaaac	caccagtgca	ggcaactgca	ccaaagatcc	cttgctgcgc	1560
ctggatgctg	ctatcagcttc	catcgacaaa	tccggttctc	ccctgggtgc	tgaacagagat	1620
cgctcggatt	cgcagctcac	caactcgac	aacacacata	ctaacctcgaa	tgaagcgcag	1680
tcccgatttc	aggacgcgca	ctatgcgac	gaagcttcca	atatgtcgaa	agcgcagatc	1740
atccagcagc	cggttaactc	tggtgtggca	aaagctaacc	aggtaccgca	gcaggttctg	1800
tctctgctgc	agggttaa					1818

<210> 51

<211> 1344

<212> DNA

<213> Escherichia coli

<400> 51

atgggcacaa	tcattaatac	caacagcctc	tcgctgatca	ctcaaaaata	tatcaacaag	60
aaccagctcg	cgtctgcgag	tcttatcgag	cgctgtctct	ctggcttgcg	tattacaagc	120
gogaagggatg	agccgcgcag	ctaggcgatt	gctaacgctt	ttaacttctaa	cattaaaggc	180
ctgactcagg	ctgcacgtaa	cgccaaagcag	ggatattctg	ctgcacagac	caactgaaggc	240
gcgctgtccg	aaatcaacaa	caactctacg	gctattctgt	aactgaacgt	tcaggctctc	300
accgggacta	actctgattc	gcattctggac	tccattcagg	acgaaatcaa	atccgctctc	360
gacgaaattg	accgcgtttc	cggtcagagc	cgcttcaacg	cgctgaacgt	ctgtggcgaaa	420
gacgggttga	tgaagattca	ggttgccgag	aatgacgggc	agacacatctc	tatcgatttg	480
cagaaaaattg	attctctcaac	gctgggattg	aaagggttct	cggtatcagg	gaacgcattta	540
aaagtctagc	atgcgataac	tacagttctc	ggtgctaatg	ctggcgatgc	cccggttaac	600
gttaaatattg	gtgcgaaacga	acagctgctg	ggcgcaatgg	ctaaacatct	cgactgaagt	660
gatacatcag	gcttggtcctc	acataacgta	caagcgcgac	atggtaaacg	gacaggaacc	720
tatgtgtgtc	aatctgggtaa	tgaattctat	tgggtctcgg	ttaatgtctg	ttgtgtgtgt	780
cgcttaataa	ccaccaaagt	tactttcaet	gatcctgcga	acggtgttac	caacgacaac	840
cagacaggtc	agcctatcaa	ggtaacgacg	aatagtctct	gcgcgctgtg	tggtctatgt	900
actatctoaag	gcaaaagatta	ccttgctggg	gcagacggtg	aggaatgcaat	tgaaaaacgt	960
ggtagcgcgtg	caacaatatga	agacacaaaa	atccaaacta	ccgatgaact	cgatgttgat	1020
gtgtctgttaa	aaacacggcg	aacagcaaca	ttttctggtg	ctgcacacaa	cgatccgctg	1080
gactcttttag	acaaagactat	ctgcgaagtt	gtactttctc	gtctctccct	cggtggcgtg	1140
caaaacccgtc	tggattctcg	ggtaaccaac	ctgaataaca	ccaccaccaa	ctgtctgtaa	1200
gcgcagctcc	gtattcagga	gcgcgactat	gcgaccgaag	gttccaaact	gtcgaaacgt	1260
cagatcatcc	agcaggcggg	taactctgtg	ctgtctaaag	ctaaccaggt	accgcagcag	1320
gttctgtctc	tgctgcaggg	ttaa				1344

<210> 52

<211> 2599

<212> DNA

<213> Escherichia coli

<400> 52

cttctcttag	ctctcttatt	gagcgtctgt	cttctgtgct	gcgtattaac	agcgcaaaaag	60
acgatgcagc	aggtcaggcg	attgctaacc	gttttaacgc	aaatattaaa	ggctgaccc	120
aggctctccg	taacgcgaat	gatggatttt	ctgttgcgca	gaccactgaa	ggtgcgctga	180
atgaaattaa	caacaacctc	cagcgtattc	gtgaactttc	tgttcagcca	actaaagta	240
ctaactctga	cagcatcttt	tcttctatcc	aggctgaaat	tactcaacgt	ctggaagaaa	300
tgacgctggt	attctgacaa	actcagttta	acggcgtgaa	agtccttgct	gaaataaatg	360
aaatgtgaat	tcaggttggt	gctaattgat	gtgaaacat	tgactctgcc	caacgattag	420
atatacaact	cagtttagtaa	cgctcggaatc	ttcatttctc	gaatgacctc	ttcttcagcc	480
cgctcgcaat	tacacgggtg	cttgataatt	cagcgtggag	tgccggcgccg	attctgtata	540
atctctccgc	cagtcattaa	taattttctc	ggcatgaaac	atatcgctga	accagtgctc	600
attcaaacat	tactcgcgaa	atcgctcggt	aaagctctca	ataaatccgt	ttctgctgtg	660
cttgcgccgc	tggtattaagc	gcaactcaac	accatgtcca	aggcccaat	gatccagtcg	720
acggcgaagt	taactccggc	cttggtcagt	tcttatcgta	gcggatagc	gtcgaaacag	780
tgcaatgctg	tccagaatac	gcgtgacctg	aaagctctgaa	atcccaaaag	caacagtacg	840

cgtcaggcat	tcccttgatga	aatcatcgac	gcaggtaaga	cacttgatcc	tgcgaccggt	900
ggaaagctgcg	tccatgacga	aatccatcga	ccaggctcaga	ttgggcgccg	ccggacggag	960
cagcggcgaga	cggtctgttg	ccagcccttt	acgagctctt	ctgcgtttta	cgcccaggcc	1020
actgagggtga	taaagccggt	acacgcgctt	atgattaa	tgaagccctt	cagggcgagc	1080
caactgctcaa	atacgcaggt	agccaaaacg	cctgcgctcc	agtgccagct	cagtgaatgc	1140
ccctgcataaa	tgcgcatcag	cagccggagc	gtgagcctca	tgcggcgagg	tcgacagcga	1200
taaacactgta	agcgtcgagg	cacgacgttg	cgaacagacg	gtcgatcac	acatacaact	1260
cacggtctcc	cgcttctggg	ctgtcgtcag	tactttcgcc	caagagccac	ctggaagcgt	1320
tctttatcca	tgcgtgcttc	ggcaagcagc	ttcttgagtc	tggtgtcttc	ttcctcaagg	1380
gacttcaggc	gcttaacttc	aggcaactcc	ataccgcat	actcttaag	cagggtgtca	1440
aacgtggcat	cggaaatggc	atgcttgcgg	cagagttcac	ggcggggtac	ccagactctcg	1500
gcttcgcggga	gaatcactgat	gatctgtctg	tcggaaaaac	gcttcttcac	ggggatgttc	1560
tcattgtgct	tatgaagaca	ttactaacat	cggggtgtac	taatcaacgg	ggagcaggtc	1620
accatcacta	tcaatctggc	aaaaattgat	gcgaaaactc	tcggcctgga	cggttttaac	1680
atcgatggcg	cgcagaaagg	aaacggcgag	gccttgattt	ctaaatttaa	agcgacaggt	1740
actgataatt	atcaaattaa	cggtaactgat	aactatactg	ttaatgtaga	tagtgagata	1800
gtacaggata	ataatggcaa	acaagtttat	gtgagtgtct	cggatgggtc	actttacgac	1860
agcgactgata	ctcaattcaa	gattgatgca	actaagcttg	cagtggtcgc	taagagattta	1920
gctcaaggata	ataagattgt	ctacgaaggt	atcgaattta	caaataccgg	cactggcgct	1980
actatttcga	caggtaatgg	taaaattaac	gcgaattgtt	atggttaagtc	ttgtgtaattc	2040
ctatccagga	ggatgtctga	tacatcaggt	actagtgcga	ccgttgcccc	taogcagcgc	2100
ctatcaaaaa	atagtgcagg	gcaattgact	gcacacaaag	ttgaaataaa	agcagcgaca	2160
ctatctgtag	ttgatctgaa	cgcgtgcaag	aaaaacaggga	gcacgttagt	tgtttaacggt	2220
gcaacttacc	atgttaagtgc	agatggtaaa	acgataacgg	agactgcttc	tggttaacac	2280
aaagtctatg	actctgagcaa	atcgaaggtg	ggtagcccca	ttctggtaaa	cgaagatgca	2340
gcataaatctg	tgcaactctc	accacaacgg	ctcgaaactc	tgacacaaag	ataggctaaa	2400
gttgacaatc	tgcggtctga	cctcggtgca	gtacaaaaac	gtttcgactc	tgccatcacc	2460
aaccttgcca	acacgtctaa	aaactgtctc	ttcgccccat	ccgctatcga	agatgctgag	2520
tacgcgacgc	aagtgtctaa	catgtctcgt	gcgcagatcc	tgcaacaacg	gggtacctct	2580
gttctggcac	aggctaacc					2599

<210> 53  
 <211> 1245  
 <212> DNA  
 <213> Escherichia coli

<400> 53						
aaacaaaaac	agtcctgcgt	gtcgacttct	atcgagcgcc	tctcttctgg	tctgcgcatt	60
aaacagcgcta	aagatgacgc	tcggggccag	gcgattgtcta	acgcgttcac	ttctaaccatc	120
aaaggtctcga	ctcaggccgc	acgtaaacgc	aacgacggta	tctctctggc	gcagaccact	180
gaaggcgac	tgctcgaat	caacaaacac	ttgcagcgtg	ttcgtgaact	gcaggttcag	240
ggcactacgc	gtactaacct	tgattctgac	ctgtcttcaa	tccaggacga	aatcaaatcc	300
cgtctcagtg	aaattgaccg	cgtatccggg	cagaactcag	tcaacggcgt	gaacgtactcg	360
gcataaagatg	gtcgcgatga	aattcaggct	ggtgcataatg	atggtcagac	aatcagcatt	420
gatttgcaga	agattgattc	ttctacttta	gggttaaatg	gttttctgt	ttccaaaatat	480
gcagtatctg	ttggtgatgc	tattactcaa	ttgcctggcg	agacggcagc	cgatgcacca	540
gtaacacatca	agtttgatga	ttcagtaaaa	actgatttaa	aactgaccca	tgcttcaggg	600
ttaaagtctgc	ataacttcaa	agatgaaaa	ggtaatttaa	ctaaccagta	tggtgtacag	660
aatggtcgga	gaactttacg	tgctacagtc	gctgccaatg	gtaattgtac	gctgaacaaa	720
gcataatgtaa	ctcagacgca	tgctgcacac	ggtattgata	ccgcaacgca	gtcaggccag	780
ttagtctcag	ttgggtcaga	ttctacgggt	acgcacaaag	cattcgtgtc	tgctccaaagt	840
aaaggtcttg	gcattgatga	cgccgccttg	aagaataaca	ctggtgatgc	tacgcgtact	900
ccacggggaa	actctgggac	acacgttgct	gcagcgtcaa	ttcactgag	tcgggcacaa	960
aactctgtac	acgtctgatg	aacggcttcc	actgaattca	cagtgctctc	aaccaacgat	1020
ccactgtagc	tgctggacaa	agctatcgca	tctgttgata	aattccgttc	tctcttgggg	1080
gcggtacaga	acgtctcgag	ctccgctgta	accaacctga	accaacccac	caccaacctg	1140
tctgacggcg	agtcctcgat	tcaggacgcc	gaactatcga	ccgaagtgtc	caacatgtcg	1200
aaagcgcaga	ttatccagca	ggcaggtaac	tcctgtcgtc	ccaaa		1245

<210> 54  
 <211> 1212  
 <212> DNA  
 <213> Escherichia coli

<400> 54  
 aacaaaaaac agtctgcgct gtcgaacttct atcgaaacgcc tctcttctgg cctgcgtatt 60  
 aacagctggca aagatgacgc tgcgggtcag gcgatagcta accgtttcac cttcaacatt 120  
 aaaggccctga ctcaggctgc gcgtaacgcc aacgcagcta tttctctggc gcagaccaca 180  
 gaaggtggcgt tgcctgaaat caacaacgac ttgcaacgtg tgcgtgagtt gaccgttcagc 240  
 gcgacgacgc gacttaactc tgattctgac ctgtctatcta ttcaggagca aatcaaatcc 300  
 cgctctggatg agattgacgc tgtttccgggt cagaccacagt tcaacggcgt gaattgactgt 360  
 gcaaaagacg cttcgtatgaa gattccaggtt ggcgcgaagt atggccagac tattagcatt 420  
 gatttacaga aaattgactc ttctacatta ggggtggaat gtcttcocgt tctcgtccaa 480  
 tcacttaacg ttggtgattc aattactcaa attacaggag ccgctggggc aaaaacctgtt 540  
 ggtgttgatt tcactgctgt tgcgaaagat ctgactactc gcacaggtaa aactgtcgat 600  
 gtttccagcc tgacgtttaca caacacctgt gatgcgaaag gggctgccac cgcaacagttc 660  
 tgcgttcaat ccggtagtag tttctactcc gcgtccattg accatgcaag tggtagaagt 720  
 acgttgaata aagccgatgt cgaatacaaa gacaccgata atggactaac gactgcagct 780  
 actcagaaga actcagctgat taaagtgcgc actgactctg acggcgccgc tgcgggatat 840  
 gtaacattcc aggttaaaaa ctacgctaca gcggctccag cggcgcttaa tgatgacatt 900  
 accgcaacag ccacagcgaa caaagtgtgt gtggaattat ctacagcaac tccgactcgc 960  
 cagttctcag gggctttctc tgcgtatcca gtcggacttt tagacaaagc cattgcacag 1020  
 gtgtgatact tccgctctcc cctcgggtgc gttcaaaaac ctgctgactc tgcgttaacc 1080  
 aacctgaaca acaccacacc caacctgtct gaagcgcagt cccgtattca ggcgcgcagc 1140  
 tatgcgacgc aagtgctcaa catgtcgaaa gcgcagatca tccagcagcg ggttaactct 1200  
 tgcgtgtcta aa 1212

<210> 55  
 <211> 1758  
 <212> DNA  
 <213> Escherichia coli

<400> 55  
 atggcacaag tcattaatac caacagcctc tgcgtgatca ctcaaaaata tatcaacaag 60  
 aaccagctctg cgtctgcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120  
 gcgaaggatg acgcccgggg tcaggcgatt gctaaccggt ttaactctaa cattaaagcg 180  
 ctgactcagg ctgcacgtaa cgcacaacgac ggtatttctg ttgcacagac cactgaagcg 240  
 gcgctgtccc aatatcaaaa caacttaacg cgtatccgtg agctgacggt tcagggtctc 300  
 accggagacta cctctgattc ggatctggac tccattcagg acgaaatcaa atcccgtctc 360  
 gacgaaatgt accgctgata cgtgcagacc cagttcaaac gcgtgaacgt actggcaaaa 420  
 gacgggttga tgaaaaattca ggttggtgag aatgacggtg aaactatca tctgcacctg 480  
 aagaatatcg attctgatac tctgggtctg aatggtttta acgtaaatgg taaaggtaac 540  
 attacacaac aagctgcaac ggttaagtgt ttaactctg ctggcgccga ggttaaacac 600  
 accacaggtc ttatgatctc gaaaaccgaa aataacctgt taactacgca tgctgattc 660  
 gataaattag ggaatggcga taaagtcaac gttggcgccg tagattatc tatcaacgct 720  
 aaatctcgtg attttactac caccaaatct actgctggta cgggtgtaga cgcgcggcg 780  
 caggtctact attcagctaa aaacgtgtat cgttagctgc cacccttca tgctgatgtg 840  
 ggttaaatctg ttaagtgttc ttacacacaa aaagatggtg ctgtttctt cgaaacggat 900  
 tcagcagcta atataccat ccggtggaag caggcatacg tagacgatgc aggcacactg 960  
 acgactaaca acgctggttag cgcagctaaa gctgatatga aagcgtgct taaagccgg 1020  
 agcgaaggtg gtgacggtgc ctctctgaac ttcaatggca ctgaataata ctgacataca 1080  
 gcaactcctg ctgaacaactc tccagtatgt ccgttaattc ctggtgggat tacttatcag 1140  
 gctcagtgca gtaaaagtgt agtatgagc gaaaccaaag ccgctccgcg gacatcttca 1200  
 attaccttta attccgctgt actgagcaaa actattgggt ttacocggcg tgaatccagt 1260  
 gatgcgtgca agtcttatgt ggatgataaa ggtggtatta ctaacgtlrg cactatatac 1320  
 gtctcttaca gcgttaacaa ggataaacgc tctgtgactg ttgcgggta tgcttccagc 1380  
 actgcataca ataagaatta tgcctcagca attggtactg cttaaatgt gaactccgcg 1440  
 ggttaaatca ctactgagac taccagtgtc ggttctgcga cgaaccaacc gcttgcctgc 1500  
 cgttcagcagc ctactcagct catcgacaaa tccgttctt cctgggtgc tatccagaag 1560  
 cgtctgattt ccgcagctcac caacctgaac aacaccacta ccaacctgtc tgaagcagc 1620  
 tcccgatttc accagccgca ctatgcgacc gaagtgctca acatgcgaa agccagatt 1680  
 atccagcagc ccggttaactc cgtgctgcca aaagccaacc aggtaccgca gcaggttctg 1740  
 tctctgctgc aggtttaa 1758

<210> 56  
 <211> 14024  
 <212> DNA  
 <213> Escherichia coli

<400> 56  
 gtaaccaaga ggcgtacgtg cataaaatctt aatgctttatc aaactattta gcattataaa 60  
 tatataaaga attctcaaat gaacaagaaa accgtttcaa taattatgcc cgtttacaat 120  
 gggggccaaa ctctaatcttc atcagtagaa tcaattatata atcaattctta tcaagatttt 180  
 gttttgtata tcaattgacga ttgtagaccc gatgataacc gatgataacc caaatttagg tggtagcaga 240  
 tacaaaaaca atcgaaaaat aagaatattg cgttaacaaga cgaatttagg tggtagcaga 300  
 agtcggaatt atcggaataga aatggccacg ggggaatata ttcttttttg tgaattgcta 360  
 gattttgtgc agcgaaaaaa attagagcgt gaatacaata atgtaataaa tgaattgcta 420  
 gatgtggtat gtcttaattt ttatgttata atgtaataaa atgtaataaa aaactacat aggggaattg 480  
 aatgctcttc atgtgataaa ttatagaaaa atgtaataaa aaactacat aggggaattg 540  
 acaggaatct ataatgcaca caaattgggt aagttttatc aaaaaaagat tggtagcaga 600  
 gattatttga ttgtggctga aataatttat tcaattatc aaaaaaagat tggtagcaga 660  
 aatctggcgt attacatgag ttcaaatatc tcaattatc aaaaaaagat tggtagcaga 720  
 aaatggacat ggagtatata tagagaacat tcaattatc aaaaaaagat tggtagcaga 780  
 tatattttat tatatgcttc aaatggagtc atgaaaaaaa tttttttatc ctattttcac 840  
 agaaaaggaga atcaaaaagtg aagtcagcgg ggggttatca tagatgatgc tatacaaaa 900  
 tttatagctc ccagttgtat ggggttatca tagatgatgc tatacaaaa 960  
 aggtattaac tagtattata attatatctc agttttttt tggtagcaga 1020  
 cgattataaa tgaagaagaaa cagcagaaaa aacttatcgt gaactgggag cttaagttaaa 1080  
 tactcgtttt cctctttgtg atcatagaaa ttgctgtcgt agttttttt cttaagaaga 1140  
 gatttctcat atttgatgat gatccagggg gggctaaact tagataagct gaagcttaag 1200  
 gactttacat tagatatatt aagtaatttg tgaattatgt tggtagcaga ttaattttc 1260  
 tttatgatga gcataaaatc aaacagagga ccatcatatt tggtagcaga ttaattttc 1320  
 cttatttttg ttatcgtctc gaattgggtg ttgtagcaga ttaattttc 1380  
 atatcctgtc aaaggataac cgtcaatcct aataaaaaa ttaattttt 1440  
 tggtagcaga ttgtgctcgt ttgttttatc taagtttagg acagacgga gaacaaaaat 1500  
 accatataaa taatatgtta aggataaata ataggttaac aatagacga ttgtagcaga 1560  
 ttccatataa tggttctgaa tctatlaaga acgattttt tccgacacca gagttagaaa 1620  
 aggaattaaa agcaataata aatagaatca aggggaataa gcatcaagac ttatttttt 1680  
 gagaacgctt acataaacaa gtaattggag acatgggagc gaatttttt tccgactacta 1740  
 cgtatggagc agaactgtta gtttttttg gttttctcgt tgaattttt atccctttt 1800  
 ggatataata acccttttat attatgattt tattgcaata cttagtgctc tggtagcaga 1860  
 ggcatttcta tttcatatct ttctctccg tattgataat gtgtactcct ctgactttat 1920  
 ggccttctt ttttgctctc ttatcacgaa atgaaaatat cagttataac ttgtagcaga 2000  
 tgcattgctg aaggggttag aaaaacttta agtagtttat cgaattttta aataaaacct 2100  
 tttagatta ttatagttga tggcggctct acagatggaa aaactgttat attagtaga 2160  
 tttagatta tgaattattc acatgttat gaaaaagat aggggataa tggtagcaga 2220  
 aataaaggag gtaattggc caaaggcagc ttaatacatc atttaaacgc cggcagtagc 2280  
 gtaattggag atatatataa aaatacctaa gagcactatt tgaattaaat tggccttttc 2340  
 gaaaatgata aaactctggg attttctctt ataacccatt caaatacagg ttattgtcat 2400  
 caaagggtga ttttccaaa gaactattca gaattatgat taaggtataa aatattgtct 2460  
 gattataagc ttattcaaga ggtgtttctt gaagggttaa gatctctatc ttgtattact 2520  
 tggggttatg taaaatatga tatgggggga gtatcttcaa aaaaaagat tttaagagat 2580  
 aaagacgttg caaaaattat gtttgaaaaa aataaaaaaa accattataa cttatttcca 2640  
 attttcaata tcaaaaattt attccttgaa cgtttaagaa gagttattgc gaaaatgcaa 2700  
 tatattgttc taactttatt ttctatgaag aatagttcac catatgataa tgaataaaa 2760  
 caaaaaataa cttaaatatt gcactttaaa aaatatgat acatcaagtg ctttaggtag 2820  
 agacacggaa aggtacagga ttatatcctt gtcgttatc tcaagttga catatgaaa 2880  
 actctcacta cttctcctta tattaaactg aagtttaact ttaactttat taggacaaga 2940  
 gagatttggt ttgtgtagta ctattaccag tctgtgact tctgtgact tttggact 3000  
 aggtatagga aatgcattaa caaacaggat cgcacattca ttgtcgtgtg gcaaaaattt 3060  
 aaagataggt cgcgaattta ttgtgtggtc caatttgcgt gctgtagtat cttttgtcat 3120  
 aactgcaata ttgtatata cttctgcat gatgtattg caactagtaa taaaggtat 3180  
 aaacgagaat tgtagtcag agttacaaca ctcaattaaa cttttgttaa tcatatttgg 3240  
 acttgaattt tattcaatg ttgtgcaaaa agtttatatt ggaatacaaa agccttatc 3300  
 aagtaataat ttatgcca tatttatatt gtatctatt attacttat ttaactgtc 3360  
 gaaactacat cgggactac cagtttttat ttgtcagcact ctgggtattc aatacatatc 3420  
 gggaactctat ttaacaata atcttattat ataaagttaa caaagattaa 3480  
 catcatagct aaaaagaaag ctccattatt gatattaaac ggttttttct tttttttt 3540  
 acagttggc actctggcaa catggagttg tgataacttt ataatactta taactatgg 3600  
 tgttacttat gttgctgttt tttagcttat acagagatta tttcaaatat ctacggctcc 3660

tctttacgatt	tataaacatcc	cggttatgggc	tgcttatgca	gatgtctcatg	cacgcaatga	3720
tactcaattt	tataaaaaga	cgctcagaagc	atcattgaaa	atagtgggtga	ttctcatcatt	3780
cttatctggcc	ttcatatag	tgctgttcggg	tagtgaagtc	gttaatatatt	ggacagagaag	3840
aaagatctcag	gtacattcgaa	catctcataat	agcttatgct	ttatggtctgt	ttatgtatgac	3900
tttttcgaat	acatttgcgaa	gcttttttaaa	tggtttgaaac	atagtttaaac	aaacaatgct	3960
tgctgttgta	acattgatata	tgatcgcaat	tcacgcaaaa	tacatcatag	ttagccatttt	4020
tggtgttaact	gttatgttgt	actgcttcact	ttttatatata	attgtaaatt	acttttataat	4080
gtatacaatgt	agtttttaaaa	aaacatatcga	tagacagtta	aatataaagat	gagtaaaaatt	4140
aaatatatata	cagttttacca	accgctcattg	acagggaaaag	aaaaagaata	tgtaaaatgaa	4200
ttgtctggact	acaactgggat	ttcatcmeta	ggaaactata	ttcagaagtt	tgaaaataaaa	4260
tgctgtggaa	aaaacatagt	aggtatatcga	gaagagagat	aaagtattgt	gggtgtcctct	4320
catctagctt	tgcttagcgtt	aggtatatcga	taacagaggag	ccaccccccct	tttcgttgtat	4380
acatatataag	catcagtttaa	tgcttatataa	aatgtatcgt	aaacaaaaaat	cactataaaaa	4440
tcagataaatg	aaacttgggca	aaatgtcttgt	ggatgacatag	gtgatattgga	acaaatttgta	4500
actaaagatg	ttatgtgtgt	ccattttatac	ggatgacatag	ggctgtgaagc	ctttgtgtct	4560
gaactgggcca	aaagttagaaa	tttgttttgta	atgtgaagatt	cgactttttag	cttttttggga	4620
aaatataaaag	gtaaaatagt	gggaacatttt	ggagatattt	cgaaatgacaa	aaacatttga	4680
aataaaaacta	ttactactcagg	tgaaaggtgga	atgggtgtac	ataggcaata	ttggcatgac	4740
gacogttgtt	tacatttttaa	agggcaacgga	ttagctgtac	ctataggatt	ttccaaactgt	4800
gttataggct	acaaattatag	gatgacaaat	attgcgctgt	ctgatatttta	taaaaaaaatt	4860
gaacaaagctg	atcatttttat	atcacgcaaaa	ctgtcaaaatg	tttttccacat	ttattggagtg	4920
atcaaacagtc	ttgttacaagt	ccacaaggaaa	agtaaaagatt	taagggaatca	cottgcagat	4980
gtctcaattc	taactaggac	cgcagaggaaa	agagagggaat	cgatgccaat	gtactgcgcat	5040
aaactctatcg	aaacaaggccc	agtttttttac	ctgtgtccaca	tttggtttggc	tttaaccttaa	5100
aaatcatcaaa	agccacctat	agctcaggagt	cttggtttggc	tttggtttggc	tttaaccttaa	5160
ttccccagcc	ttcaggaatga	gcaaggttatt	ttatcttatga	aaatttgcgtt	gaatttcagat	5220
agtgataaaat	agctcaaaaat	attgcaaaagg	ttatcttatga	aaatttgcgtt	gaatttcagat	5280
ggatttttaag	agtgggggcgg	tggaatttgat	ttattttaa	aaatttgcgtt	gaatttcagat	5340
acgaataccag	aaatatgtat	cgatatcttt	ttaccggagaa	atgataataca	ttcttttata	5400
agagaaaaag	catttcccttt	taaaaagtata	ttataaagca	ttataaagag	ggaaagggcct	5460
cgatgtgatt	catctaaatg	atttaactgag	caaatcactga	gagatgcctt	ttccaaaaat	5520
aaatatagaga	cgatctctac	ctttatcaaa	agtaaaagct	ctgccttttta	ttcatatttt	5580
gatagtagat	atttgtgatg	tattcttccct	tgcatgtgtgt	ttcctctggg	aaattttgaat	5640
aaaaaagcat	ggatagggtta	tatttatgac	tttcaacscg	gttaactatcc	ttcatttttt	5700
agtagcagag	agatagatca	aagggaatgtg	tttttttaaa	gatgctcaca	ttgcgctaac	5760
aaatatatgt	ttatgacaca	tttcagtkatt	accgtagcca	ataaatatgt	tggaagattat	5820
ctctgcaaac	tacattctct	ttcatattag	ccatgcccct	aattaaaatg	gttcgctgat	5880
tactctggta	ataattgcaaa	ataataatgt	gacaaagagt	attttataat	ttgcaaatcaa	5940
ttttgcaagc	ataaaagatca	tgcaactgct	tttagggcct	ttaaaaattta	tactgaaat	6000
aaactcgtatg	tttatttagt	attgcacggga	gctactcaag	attatcgtat	ccctggatata	6060
tttaatgaat	tgatggtttt	ggcaaaaaag	ctcggaattg	aatcgaataat	taagatatta	6120
gggcataatac	ctaaactcga	acaaatgaa	ttaatcaaaa	attgcattgc	agatatacaca	6180
ccaaaccttat	ttgaaggcgg	gctcggaggg	ggggtaacat	ttgacgctat	tgcatattgg	6240
aaaaaaggta	tactactcga	catagatgct	ataaaagaag	ttaattgcgg	tgatgtatgt	6300
ttcttttcag	caaaaaacca	ttaactcata	aatgacgcga	tggttaaaagc	tgatgaactct	6360
aaaaattttt	atgaacctac	aactctgata	gaatttgggt	tcaaaagacg	caatgcgctg	6420
cgagattttc	tttttagatg	tgtagaaca	gaaattgaa	cccgatctta	atacatatcaa	6480
gaggtatata	atgactaaag	tcgctcttat	tacagggtga	actggacaac	atggattctta	6540
cttagctgag	tttttgctgt	ataaagggtta	tgaaagtctat	ggatcaaac	gcgcagccct	6600
attcttttaa	acagaacgca	tagaccatata	ttatcaagat	ccacatggtt	ctaacccaaa	6660
ttttcaactg	cactatggag	attgactgta	ttcatctaac	ctcaactgaa	tttcaaaaga	6720
ggatcagccga	tgaaagrat	ataatttagc	tgctatgagt	caagtacagc	ttctttttga	6780
gtctccagaa	tatacagccg	atgtcgatgt	aatttggtaca	ttacgtttac	tggaagcaat	6840
tcgcttttta	ggattggaaa	acaaaacgcg	ttctatcaca	gcttcaacct	cgagatata	6900
tggaacttgt	caggaatacc	ctcaaaaaga	atccaccctt	ttttatctct	gttccccctta	6960
tgcaagtgtga	aaacttttac	catatttgat	cacggttaaat	tatcgagagt	catatggat	7020
ttatgcagt	aattggtatat	tgttcaatca	tgaaatctcca	cgccgtggag	aaacgtttgt	7080
aaacaaggaaa	attactcgcg	gaacttgcaaa	tattgcacaa	ggcttggaat	ctgattttga	7140
tttagggaaat	atggatttgt	tacagagattg	gggacatgca	aaagattatg	ttagaatcga	7200
atgggttgatg	ttacaacagg	agcaaccgga	agattttgtg	attgcaacag	tgactaata	7260
ctcagtcctg	cagtttgtcg	aaatggcagc	agcacaactt	ggatataaga	tgagctttgt	7320
tggttaaaaga	atcgaagaaa	aggcatttgt	agattcgtgt	gaaggacagg	agctctcagg	7380
tgtagaaacca	ggtagatgca	ttgttgtcgt	tgatctcgt	tatttccagc	cagctgaagt	7440
tgatatcttg	ctggagagct	cagacaaaagc	taactctcaa	cttggttggga	cagcagaaa	7500
tactcttttgt	gaaatgattt	ctgaaatggt	tgccaaaagat	cttgaagccg	ctaaaaaaca	7560
ttctcttttgt	aaatcgcagt	ttttttctgt	agacttttagt	ctggaatgat	gatgataaag	7620
caacgtattt	ttattgtctg	tcaccaaggaa	atggtttgagt	cagctattac	ccgacgccc	7680

aaacaacgctg	atgatgttga	gttgggttta	cgtaactcggg	atgaattgaa	cttggttgatt	7740
agtagacgctg	tttttgattt	ttttttcttca	cagaaaatcg	accagggttta	tttggcagca	7800
gcocaaagctg	gaggttatatt	agctaacagat	tcttatcctg	ccgatttttt	atatgagatt	7860
ataatgatag	aggcggaatg	cattctatcgt	gccacacaaa	ataatgtaa	taaacctgtt	7920
ttcctcggtt	cgctggttat	ttatcctaag	gtacacacac	aaccgattat	ggaagacgaa	7980
ttattacaag	ggaaactctga	gccacaacat	gaaccttatg	ctatcgcaaa	aattgcaggt	8040
attaaattat	gtgaattctta	taacocgtcag	tttggggcgt	attacocgttc	agtaattgcc	8100
accaattctt	atggtccaaa	tgacaatttt	catcccaagt	attctcatgt	gattcccgctg	8160
cttttgcgct	gctttcatga	tgctgtggaa	acaactcttc	cgaattgtgt	tggttgggga	8220
agtggttactc	caaaagcgtga	attcttatcat	gtagatagta	tggtctctgc	aagcatttat	8280
gtcatggaga	tgccatcaga	tatatggcaa	aaaaatgta	aagtaaatgt	gtctcatatc	8340
aatattggaa	caggtatttga	ctgcacgatt	tgtgagcttg	cggaaacat	agcaaaaagt	8400
gtggtttata	aagggcatac	tacgttcgat	acacaaagcg	ccgattggagc	ccctccgaaa	8460
ctacttgatg	taacgcttct	tcatcaacta	gggttgaatt	ataaattac	ctctcaacag	8520
ggctcttgaaa	atcacatacaa	ctggtttctt	gaaaaccacac	ttcaatatcg	ggggtaataa	8580
gtttttttaca	ttcccaacag	tttgccacaa	ttgttaaggct	taactctctt	attctctatg	8640
atttgattgt	ggaaaaacgag	tttggcgaaa	ttttgctagg	aaaaacgaac	aaccgcccggt	8700
cacagggctga	ttggttcgct	cctggttggt	gggtgttgaa	agatgaaaaa	ttgcagacag	8760
cccttgaacg	attgacagaa	attgaactga	gaattcgttt	ctctctctct	gtgggttaagt	8820
tttatgggat	ctggcagcac	ttctacgaag	acaatagtat	ggggggagac	ttttcaacg	8880
attatatagt	tatatgattc	cttcttaaat	tacaacaaa	catcttgaaa	ttacccgaag	8940
cacaaacata	tgctttttgc	tggtctatcg	gagcaaaagt	gataaatgat	gacgattgct	9000
attataattg	tcggcgaat	tttaacata	aaacaaacga	tgccattggc	ttagatataa	9060
aggatataat	atgtctctgat	cgccaaata	tgctgtatgt	tttggccggt	gtcacggcgag	9120
tcgtcttcgt	coactttctc	gtgaactata	tcocaaagcag	tttttaaac	ttctctggta	9180
taaacaccttg	ttacaaaacga	ctttgctacg	accttcacgc	ctatcargtc	aaaaaagctt	9240
agtgataaca	aatgaaacag	atcgctttgt	ttgggtcgaa	cagttaaagg	aaataaaata	9300
atttaattgt	aatatttttc	tagaaccatg	cgggcgaaat	ccattgcttc	aatagcgat	9360
attctgctgt	catcgcttaa	aacgtaatcc	tcaggaagat	gctattaaaa	ttagtctctg	9420
ggcagaccac	gttatagcta	aagaaagctt	ttctctgtat	ccagaataatg	ctgaaactgg	9480
catcgctcaat	caaggtaaaa	ttgtaacggt	ttgaattata	caaggcgctt	aaaaactagg	9540
ttatgggat	attgagagag	gtgaactatc	tgtacgcgtt	gacactcgtt	aaattgtat	9600
ttttttttat	tgaaataagt	ttgtcgaaaa	gcctaatcgt	ccagcccgag	aatgttatat	9660
gacttcggtg	aacctcaatt	ggatagtggt	aatattctat	tttaagcgct	ctgtttatct	9720
tgaggaaatg	agaaaattta	gaactgacat	ttacaagtgt	ttgtgaacag	tggtcctcat	9780
ctctacacatt	gatctagatt	ttattcgatt	atcaaaaagaa	aaacttcaag	atrgtctcgc	9840
tgaattcatt	gttttttctg	taactggaaa	aaacagaaaa	ttgtgtgtat	gccctgttga	9900
tatttggttgg	atgtcagctg	gatcttggtc	atcgcttatgg	gacattagtc	taaatatcgaa	9960
aacaggagat	gtagtaaaag	gtgatataat	aaactatgat	actaagaata	attatatcca	10020
ctctgagltca	cgcttggttag	ccgccaattgg	aattgaagat	atggttatcg	tgcaaaactaa	10080
agatgcgctt	cgctgtctta	aaaagatgta	tgtaacagcat	gtaaaaaaaa	tgactcgaatt	10140
cgcttaaatgt	cagcaacgta	cagagtatat	tagtcatcgt	gaagttttcc	gaccattgggg	10200
aaaatttgat	tcgatttgacc	aaggtgagcg	atacaaaagtc	aagaaaaatta	ttgtgaaacct	10260
tggtgaaggga	aaagttaacc	ttggcgataa	ccatcgttct	gaacattgta	tcgtgtcttcc	10320
atacattccc	cttggcgacg	cgataatgct	aaactaaacta	gtcacctgca	atgataatgat	10380
ttatgaaagt	agtttcagggg	attattttggg	tgagaatccg	ggcataatcc	ctcttaattct	10440
ttaacaaacat	gaagatttaac	atatgaatac	agagatgat	attataagac	agaaagaacg	10500
cgggaaattta	ggcgaagaac	tgaattgaaga	tattgctcgg	cgcatctggc	gtgcctatgt	10560
cgaaattcttc	aaacgcgaaaa	ccattgtttt	aggcgtgat	ctcgccctca	ccagcgaagc	10620
gttaaaactg	ggcgtctcga	aaggtttaca	ggatgcggcg	gtcgatgtgc	tggtatctcg	10680
tatgtccgcg	accgaagaga	ttcattttgc	caactgtccat	ctcggaagtgc	gtgcggcgat	10740
cgaagtctacc	gccacgcata	accocgatga	ttacaacgcg	atgaagctcg	tgcccgaaagt	10800
ggctctggccg	atcagcgtgt	ataccggact	gcgcgatgtr	cagcgtctgc	cagaagccaa	10860
tgacttcccc	ctgtctgatg	aaaccaaagc	tggtcgctat	cagcaaatca	atctcggtga	10920
cgcttactgt	gatacctcgt	tcggttatat	acaacgtaaa	aaactcacgc	cgctcgaact	10980
ggtgatcctac	tcagggaaacg	gcgcagcggg	tcggttggtg	gacgcacgtg	aagcccgatt	11040
taagaccctc	ggcgacccgg	tggaaattat	caaaagtacac	aaacacgcgg	acggcaattt	11100
ccccaacggt	attctcaacc	cgctgctgcc	ggaaatccgc	gacgaacccc	gtaatcggtg	11160
catcaaacac	ggcccgagata	tgggcattgc	ctttgatggc	gattttgacc	gtcttttcc	11220
gtttgacgaa	aaagggcagt	ttatcgaggg	ctactacart	gtcgccgtcg	tggcgaacac	11280
gtctctcgaa	aaaaatcccg	gcgcgaagat	catccacgat	ccacgtctct	cttggaacac	11340
ctgtgatgtg	tgtaactccg	caggcggcac	cccggttaatg	tcgaaaacgc	gacacgcctt	11400
tattaaagaa	cgtagtcgca	aggaagaacg	catctacggt	ggcgaaatga	gcgctcacca	11460
ttacttccgt	gatttctcgt	actgcgcag	cgccatgtrc	ccgtgtcgct	tggtgcgca	11520
actcgtgtgc	ctgaagaagaa	aaacgcctgg	cgaaatgggt	gcgacccgga	tggtggcgtt	11580
tcocgcgaagc	ggtgagatca	acagcaaat	ggcgcaaccc	gttgaggcaa	ttatctcgct	11640

ggaacagcat ttagccgcg aggcgcgtgc ggtggatcgc accgatggca tcagcatgac 11760  
 cttgcccagc tggcgcttta accctgcgtc ctccaacacc gaaccggttg tggcggtgaa 11820  
 tgtggaaatca cgcgggtgat taaagctaat ggaagaaata actaaagctc tctcttaatt 11880  
 gctaagttag tgatatttta cattaatcat taagcgtatt taagattata taaagttaat 11940  
 gttattgcgg tatatgatga atatgtgggc ttttttatgt ataacgacta taccgcactt 12000  
 ttacctagga aagattaat agaataaag ttttgactgt accaatttgc atttcagctc 12060  
 acgatggaga cgttcccttg cttaagacat tttttcatcg cttatgtaat aacaaagtgt 12120  
 ccttatataa aagggagaac aaatggaaac ttaaaaatcg ttacaaatcg gatttttatt 12180  
 atccctgttt acgatattat agccaaagtt gtatcctgca tcagtcctgc aatatttcac 12240  
 gagtgccttg ttaactgaat acatgtctgc cattttccag atgataacga cgtcatcgca 12300  
 attgaggtta aaacactctg gcacacttat gacaaagatg gtccgacagg agtgggtcat 12360  
 gtcatatagt cgtttcagca atgcacagtc tggctcctcg atagatcaag accgatgaga 12420  
 aacctaatgc gttccacagt attcatgaac tttctaaat gatgggtatt aaaggaaaaa 12480  
 taatcataac tgalgcgatg gcttgcaga agaatattgc agagaagata taaaaaacga 12540  
 gatgtgatta ttatttgcgt ttaataaact caaaacatga cagttacgca atagtgaia 12600  
 agatatttac gctgaagaaa gtccgtcttc atattgtttg agatgctcct gatgagctta 12660  
 agaggacagg cagagacgat aaagggtgc agaatttat aatggcagtc gatctctct 12720  
 ttgatattac gtttgaatgg aaagggtgc aaatgaatc caaatattat atagatctg 12780  
 caataataac agacaaaag ttgcaaatcg ttaactgaat tcaactgcgc atggagaga 12840  
 ctgctttaac cgcagagaag ttgcgcacag tggtaatgaa tgaatctgac tataataaa gaaggcgagt 12900  
 agttgcaacg tagcctgatg tttgtatgaa catcgtcatt aatatcgac aattgataatg 12960  
 tgcatctgaa tgattttcta gaatgcggca tttgtatggt gaaaagcagt catggacaga aacttctctg 13020  
 tattcaaggc aggatattca tgaagatgac aaactttgat tggtttgat aggatatttc 13080  
 cgtcaggcat tgcagcgtgc gggctttcat atttttggt ctggaagtta tggtagaaga 13140  
 ttgtggagtg ggaataatgaa ttgtatggt acaaataaag caagaatgtg gttctgacta gttcttggtt 13200  
 acaataccca ttctaaatca aggaagaaaa gttaatggt ttgaagtgtc ttcaaccaac 13260  
 ttgtggagtg atgttttggc tttaaaaaag tatttttaa ttgactatg ttgaacaaat aactataaaa 13320  
 tgctttctaa aagccctcta tttcaaaaag gttcaatata ttaacacggg ttgaacaaat aactataaaa 13380  
 atacagcaga gaagctctga gttcaatata ttaacacggg ttgaacaaat aactataaaa 13440  
 catcaaaatc gcgttgttta tgatcatact atgatatgta gtggcgctat ttgctctccc 13500  
 ttgtttacaa tatctactca aatcaggagc ttgctctccc ttgctctccc 13560  
 taacgtgcac atgatgttga agacaatgca tatataagct cgggtgcagt aataaagac 13620  
 aatggatcga tttgtatgta atgcgcact tatattggc gcggagacaa ttatagttat gggggctgtt 13680  
 ggtgttctca atgcgcact tatattggc gcggagacaa ttatagttat gggggctgtt 13740  
 gctactaaaa gtgttctgc cgttatatc gtgtgcgaa atccagcaag agaaatgaaa 13800  
 agatgcgcaa catctattta atgggaatgc gaaaacagct tccaaatggg actaatgttt 13860  
 aaaaatata taatttctgt aatttactaa attatgcctt ctttttaagc tatcctttac 13920  
 ttagtattta ctgatacagc atgaaattta taactactgt atacattttt atacgttatt 13980  
 caagccgat atctagcgtt aacctctgac agggatataac aatg 14024

<210> 57  
 <211> 1758  
 <212> DNA  
 <213> Escherichia coli

<400> 57  
 atggccaaag tctaataatc caacagcctc tgcgtgatca ctcaaaataa tatcaacaag 60  
 aacagctcgt cgtctgcgag ttctatcgag cgtctgtctt ctggcttgcc tattacaacg 120  
 gcgaaggatg acgccgcagg tcaggcgatt gctaaccggt ttacttctaa cattaaaggc 180  
 ctgactcagg cgcgcggttaa gcgaacagg ggtatttctg ttgcgcagac caccgaaggc 240  
 cgtcgtctcg aaatccaaac caacttacag cgtattcgtg aactgacggt tcaggcgact 300  
 cagcgagga ttactcgattc tgaactggac tccactccag acgaatatca atctcgtctt 360  
 gatgaaattg acccggtatc cggccagacc cagtccaagc gogtgaaagt gctggcgaaa 420  
 gcgcttcaa tgaaaattca gtttggcg aatgacggc aaacacac gatcgacctg 480  
 aaaaataatg attctgatac tctgggtctg aatggcttta acgtaaatgg taaagtactt 540  
 attaccaaca aagctgcaac ggttaagtgt ttaacttctg ctggcgcgaa gttaaacacc 600  
 acgacaggtc tttatgatct gaaaacggaa aataccttgt taactaccca tgcgtcattc 660  
 gataaattag ggaatggcga taaaatcaac ttggtggcgt tagattatca ttacaacgct 720  
 aaatctgggt attttactac caactaaatc actgctggta cgggtgtaga cgcgcggcgt 780  
 caggctcgtg attcagcttc aaaaagtgat cgtttagctg ccacccctca atctcgtctt 840  
 gtttaaatct ttaattggtt ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900  
 tcagcaggtta atataccat cgggtggaagc caggcatatg tagacgatgc aggcacactg 960  
 acgactaaac acgctggtag cgcagctaaa cgtgatatga aagcgtgctt caaagcagcg 1020  
 acgcaaggta gtagcggtc ctctctgaca ttcaatggca cagaataac catcgcaaaa 1080  
 gcaactcctg cgcacacacac tccagtatgt ccgttaatcc ctggtgggat tacttatcag 1140  
 ctgcacagt gtaaaagtgt agtatggagc gaaacaaaag cggctgcgcg gacattctca 1200  
 attaccttta attccgggtt actgagcaaa actattgggt ttacgcgggg tgaatccagt 1260

gatgctgcga	agtcttatgt	ggatgataaa	ggtggtatca	ctaacgttgc	cgactataca	1320
gtctcttata	cggttaacaa	ggataacggc	tgctggaact	tgcccggtga	tgcttaacgc	1380
actgatacca	ataaagatta	tgctccagca	attggtactg	ctgtaaatgt	gaactccggc	1440
ggtaaaaaat	ctactgagac	taccagtgct	ggttctgcga	cgaccaaccc	gctgtgcgcc	1500
ctggacgacg	caatcagctc	catcgacaaa	tccgttctct	ccctgggtgc	tatccagagc	1560
cgctctggat	ccgcagctac	caactggaac	aacaccacta	ccaactgtgc	cgaaaggcga	1620
tcccgatttc	aggagccgca	ctatgcgacc	gaagtgtcca	acatgtcgaa	agcgagatgc	1680
attcagcagg	ccggtaaact	cgctgctgca	aaagctaacc	aggtaccgca	gcaggttctg	1740
tctctgctgc	agggttaa					1758

<210> 58  
 <211> 1758  
 <212> DNA  
 <213> Escherichia coli

<400> 58						
atggcacaag	tcattaatac	caacagcttc	tcgctgatca	ctcaaaaata	tatcaacaag	60
aaccagctcg	cgctgtcgag	ttctatcgag	cgctgtctct	ctggctgtgcg	tattaacagc	120
cggaaggatg	acgcagcggg	tcaggcgatt	gctaaccggt	ttacttctaa	cattaaaggc	180
ctgactcagg	ctgacagtaa	cgccaacgac	ggatattctg	ttgcgcagac	caccgaaagg	240
cgctgtctcg	aaatcaacaa	caacttacag	cgattctcgt	aactgacggt	tcaggccact	300
acagggaacta	actctgattc	tgacctggac	tcctatccag	acgaaatcaa	attctgtctt	360
gatgaatttg	acgcggtatc	cgccgagacc	cagttccaac	cgctgaaacg	gctggcgaaa	420
gacgggtcaa	tgaaaattca	gggtgggtgc	aatgacggcg	aaacatcac	gatcgacctg	480
aaaaaaatcg	attctgatac	tgtgggtctg	aatggcttta	acgttaattg	taagtactac	540
atttacaaca	aagctgcaac	ggtaagtgat	ctaacctctg	ctggcgcgaa	gttaaacacc	600
acgacaggtc	tttatgatct	gaaaaccgaa	aataacctgt	taactaccga	tgctgcatct	660
gataaattag	ggaaatggca	taaaagtcaa	gttggcggcg	tagattatca	ttacaacgct	720
aaatctgggt	attttactac	caactaaatc	acgtctggga	cggggtgtaa	cgccgcggcg	780
caggctcgtg	attcagcttc	aaaaacgtgt	cggttagctg	ccacccctca	tgctgatgtg	840
ggtaaatctg	ttaatgggtc	ttacaccaca	aaagatggta	ctgtttctct	cgaaaacgga	900
tcagcaggta	ttatcaccat	cggtggaagc	caagcacaac	tagacgatgc	agggcaactg	960
acgactaaac	acgctggtag	cgagctataa	gctgatataa	aagcgctgct	caaaagcagg	1020
acgcaaggta	gtgacggtag	ctctctgaca	ttcaatggca	cagaatatac	catcgcaaaa	1080
cgaaactctg	cgacaaccac	tcagtagctg	cggttaacct	ctgggtggat	tacttatcag	1140
gctacagtga	gtaaaagtgt	actatgtggg	gaaaccacaa	ctggctcggc	gacatcttca	1200
attaccttta	attcgggtgt	actgagagac	agtattggag	ttacgcgggg	tgaaatccagt	1260
gatgctgcga	agtccttatgt	ggatgacaaa	ggtggtatca	ctaacgttgc	cgactatata	1320
gtctctttaca	cggttaacaa	ggataaacgc	ttctgtgact	tgcccgggta	tgtctcaagg	1380
actgataacca	ataaagatta	tgctccagca	attggcactg	ctgtaaatgt	gaactccggc	1440
ggtaaaatca	ctactgagac	taccagtgct	ggttctgcga	cgaccaaccc	gctgtgcgcc	1500
ctggacgacg	caatcagctc	catcgacaaa	tccgttctct	ccctgggtgc	tatccagaaac	1560
ctctcggtat	cgccggtcac	caactctgac	aacaccacta	ccaactgtgc	gaaagcgacg	1620
tccgtatttc	aggacgcgca	ctatgcgacc	gaagtgtcca	acatgtcgaa	agcgagatgc	1680
atccagcagg	ccggttaact	cgctgctgca	aaagctaacc	aggtaccgca	gcaggttctg	1740
tctctgctgc	agggttaa					1758

<210> 59  
 <211> 1758  
 <212> DNA  
 <213> Escherichia coli

<400> 59						
atggcacaag	tcattaatac	caacagcttc	tcgctgatca	ctcaaaaata	tatcaacaag	60
aaccagctcg	cgctgtcgag	ttctatcgag	cgctgtctct	ctggctgtgcg	tattaacagc	120
cggaaggatg	acgcagcggg	tcaggcgatt	gctaaccggt	ttacttctaa	cattaaaggc	180
ctgactcagg	ctgacagtaa	cgccaacgac	ggatattctg	ttgcgcagac	catgaaaggc	240
cgctgtctcg	aaatcaacaa	caacttacag	cgattctcgt	agctgacggt	tcaggctctt	300
acagggaacta	actctgattc	ggatctggac	tcctatccag	acgaaatcaa	attccgtctt	360
gatgaatttg	acgcggtatc	cggtcagacc	cagttccaac	cgctgaaacg	atcggaacaa	420
gacgggtcaa	tgaaaattca	gggtgggtgc	aatgacggcg	aaacatcac	tatcgacctg	480
aaaaaaatcg	attctgatac	tgtgggtctg	aatggtttta	acgttaattg	taagtactac	540
atttacaaca	aagctgcaac	ggtaagtgat	ctaacctctg	ctggcgcgaa	gttaaacacc	600
acgacaggtc	tttatgatct	gaaaaccgaa	aataacctgt	taactaccga	tgctgcatct	660
gataaattag	ggaaatggca	taaaagtcaa	gttggcggcg	tagattatca	ttacaacgct	720
aaatctgggt	attttactac	caccaaatct	actgctggta	cggtgtgaga	cgccgcggcg	780



caggctactg	attcagctaa	aaaacgtgat	gcgttagctg	ccacccttca	tgctgatgtg	840
ggtaaatctg	ataatcggtc	ttacaccaca	aaagatggta	ctgttttctt	cgaaacggat	900
tcagcaggta	ttatcaccat	cggtggaagc	caggcatacg	tagacgatgc	agggcaactg	960
acgactaaca	acgctcgtag	cgacgctaaa	gctgatatga	aagcgtgtct	taaaacccgc	1020
agcgaaaggta	gtgacggtgc	ttctctgaca	ttcaatggca	ctgaatatatc	tatcgcaaaa	1080
cgaaactcctg	cgacaaacctc	tccagtagct	ccgttaactc	ctgggtggat	tacttatcag	1140
gctacagtaga	gtaaaagatg	agtattgagc	gaaacccaag	cggtcgccgc	gacatcttca	1200
attacccttta	attccggtgt	actgagcaaa	actattgggt	ttaccccggt	tgaatccagt	1260
gatgctgcga	agttcttatgt	ggatgataaa	ggtgggtatc	ctaactgtgc	cgactatata	1320
ttctctttaca	cggttaacaa	ggataacggc	ttctgtactg	ttgcgggata	tgcttcacgt	1380
actgatacca	ataaagatta	tgctccagca	attggtactg	ctgtaaatgt	gaactccgcg	1440
ggtaaaatac	ctactcgagc	taccagtgct	ggttctgcga	cgaccaaccc	tctgtgtccc	1500
ctggacgagc	ctatcagctc	ctccgataaa	ttccgttctt	ccctgggtgc	tatccagaac	1560
cgctctggatt	ccgcagctcac	caacctgaac	aaacaccta	acactctgtc	tgaagcgagc	1620
tcccgatttc	aggaagccga	ctatgcgacc	gaagtgtcca	acatgtcgaa	agcgagattt	1680
attccacagg	cggttaactc	cgtgctggca	aaagccaacc	aggtaccgca	gcaggttctg	1740
tctctgctgc	agggttaa					1758

<210> 60

<211> 1758

<212> DNA

<213> Escherichia coli

<400> 60

atggccacaag	tcattaatac	caacagcctc	tcgctgatca	ctcaaaaata	tatcaacaag	60
aaccagctctg	cgctgtcgag	ttctatcgag	cgctgtctct	ctggcttgccg	tattaaacagc	120
gcgaagagtag	acgcgcgagg	tcaggcgatt	gctaaacggt	ttacttctaa	cattaaaggc	180
ctgactcagg	cgcccgctaa	cgccaaacgc	ggatattctg	ttggccagac	caccgaagcg	240
cgctgtccgc	aaatcaacaa	caactctacg	cgattctgtg	aactgacggt	tcaggccact	300
acagggaactc	atcccgatcc	cggtctggac	tccatccagg	acgaatctaa	atctctgtct	360
gatgaaattag	acgcggatct	cgcccgagac	cagttcacgc	gggtgaaagt	gctggcgaaa	420
gcagggttcaa	tgaaaattca	ggttgggtgc	aattgacggc	aacacataac	gatcgacctc	480
aaaaaaatag	attccgctac	ttctgggtctg	aattggctta	acgttaattg	tgaaggtact	540
atacccaaca	aagctgcgac	ggtaagtgat	ttactctctg	ctggccgcaa	gtttaaacac	600
acgacaggtc	tttatgatct	gaaaacggaa	ataactctgt	tactacgca	tgctgaattc	660
gataaaattag	ggaaatggcga	taaaagtaca	gttggccgct	tagattatac	ttacacacgt	720
aaaatctggtg	attttactac	cactaaatct	actgctggta	cggtgtgaga	cgcccgccgc	780
caggctgctg	attcagcttc	aaaacgtgat	gcgttagctg	ccacccttca	tgctgatgtg	840
ggtaaatctg	ttatgggttc	ttacaccaca	aaagatggta	ctgttttctt	cgaaacggat	900
tcagcaggta	ataacacat	cggtggaagc	caggcatacg	tagacgatgc	agggcaactg	960
acgactaaca	acgctggtag	cgacgctaaa	gctgatata	aagcgtctgc	caaagcagcg	1020
agcggaagta	gtgacggtgc	ctctctgaca	ttcaatggca	cagaatatatc	catcgcaaaa	1080
gctaaactctg	gcacaaacac	tccagtagct	ccgttaactc	ctggtgggat	tacttatcag	1140
gctacagctg	gtaaaagatg	agtattgagc	gaaacccaag	cggtcgccgc	gacatcttca	1200
attacccttta	attccggtgt	actgagcaaa	actattgggt	ttaccccggt	tgaatccagt	1260
agttcttatgt	ggatgataaa	gggtggtatc	ctaactgtgc	cgactatata	1320	
ctctctttaca	cggttaacaa	ggataacggc	ttctgtactg	ttgcgggata	tgcttcacgt	1380
actgatacca	ataaagatta	tgctccagca	attggtactg	ctgtaaatgt	gaactccgcg	1440
ggtaaaatac	ctactcgagc	taccagtgct	ggttctgcga	cgaccaaccc	tctgtgtccc	1500
ctggacgagc	caatcagctc	catcgacaaa	ttccgttctt	ccctgggtgc	tatccagaac	1560
cgctctggatt	ccgcagctcac	caacctgaac	aaacaccta	acactctgtc	cgaaagcgag	1620
tcccgatttc	aggaagccga	ctatgcgacc	gaagtgtcca	acatgtcgaa	agcgagattt	1680
attccacagg	cggttaactc	cgtgctggca	aaagccaacc	aggtaccgca	gcaggttctg	1740
tctctgctgc	agggttaa					1758

<210> 61

<211> 1758

<212> DNA

<213> Escherichia coli

<400> 61

atggccacaag	tcattaatac	caacagcctc	tcgctgatca	ctcaaaaata	tatcaacaag	60
aaccagctctg	cgctgtcgag	ttctatcgag	cgctgtctct	ctggcttgccg	tattaaacagc	120
gcgaagagtag	acgcgcgagg	tcaggcgatt	gctaaacggt	ttacttctaa	cattaaaggc	180
ctgactcagg	ctgcacgtaa	cgccaaacgc	ggatattctg	ttggccagac	caccgaagcg	240
cgctgtccgc	aaatcaacaa	caactctacg	cgattctgtg	aactgacggt	tcaggccact	300

acagggaacta	actccgattc	tgacctggac	tccatccagg	acgaatacaa	atctcgtctt	360
gatgaaatttg	accgcgattc	cgcccgagac	cagttccaacg	cggtgaacgt	gctggcgaaa	420
gaacgggtcaa	tgaaaaattca	ggttgggtcg	aatgacggcg	aaaccatcac	gatcgacctg	480
aaaaaaatcgt	attctcgatcc	tggtgggtctg	aatggctttca	acgtaaatgg	taaaggtact	540
attaccaca	aagctgcaac	tgtaagtgtat	ttaaactctcg	ctggcgcgaa	gttaaacacc	600
acgacaggctc	tttatgatct	gaaaaccgaa	aatcaactctg	taactacgga	gtctgcattc	660
gataaaattag	ggaatggcgga	taaaagtca	gttggcgcgga	tagattatac	ttacaacgct	720
aaatctgggtg	attttactac	caataaaatct	actgctaggta	cggtgttaga	cgccgcgccg	780
caggctgctg	attcaagcttc	aaaaacgtgat	cggttagctg	ccacccttca	tgctgatgtg	840
ggtaaaattctg	ttaatggcttc	ttacaccaca	aaagataggta	ctgtttcttt	cgaaacggat	900
tcagcaggta	ataaccacat	cggtggaaag	caggcatacg	tagacgatgc	aggcaacttg	960
acgactaaca	acgtcggtag	cgcgactaaa	gctgatatga	aagcgctgct	caaaagcagc	1020
agcgaaaggta	gtgacggctgc	ctctctgaca	ttcaatggca	cagaataac	ctctcgaaaa	1080
cgaaactcctg	cgacaaccac	tcacgtagct	ccgttaattcc	ctgggtggat	taattatcag	1140
gctacagtgga	ttaagatgt	agttattgagc	gaaaccaaaag	cggtctccgc	gacatcttca	1200
attacacctta	attcccgctgt	actgagcaaa	actattgggt	ttaccgcggg	tgaatccagt	1260
gatgctgcga	agtccttatgt	ggatgataaa	gggtgataca	ctaactgtgc	cgactataca	1320
gtctcttaca	gcgttaacaa	ctcgaaccgc	tcgtctgactg	ctgtgaaatg	gaactccgcg	1380
actgatacca	ataaagatta	tgctccagca	attggcactg	cgacaaaccc	gactgtcgtc	1440
ggtaaaatca	ctactgagac	taccagtgt	gggtctgcga	ccctggggctg	tatccagaac	1500
ctggacgagc	caatcagctc	catcgacaaa	ttccgttctt	ccaaactctg	cgaaagcagc	1560
ctcctggattt	ccgctggctac	caacctgaac	aacaccacta	acatgtcgaa	agcgagactg	1620
tcocgtattc	aggacgcgga	ctatgcgacc	gaagtgtcca	aggtaccgca	gcaggtttctg	1680
atccagcagg	ccggttaactc	cggtgctggca	aaagctaacc			1740
tctctgctgc	agggttaa					1758

<210> 62

<211> 1758

<212> DNA

<213> Escherichia coli

<400> 62

atggcacaag	tcattaatac	caacagcctc	tcgctgatca	ctcaaaataa	tatcaacaag	60
aaccagctctg	cgctgtcgag	ttctatcgag	cgctgtctct	ctggcttcgc	tattaaacagc	120
gcgaaggatg	acgcgcgcgg	tcaggcgatt	gctaaccggt	ttactttcaa	ctaaaaaggc	180
ctgactcagg	ctgcacgtaa	cgccaaacgac	ggtattttctg	ttgcacagac	cactgaaggc	240
gcgctgtccg	aaatcaacaa	caacttacag	cgatatccgtg	agctgacggt	tcaggctctct	300
acggggacta	actctgattc	ggatctggac	tcatttcagg	acgaatacaa	attcccgcttc	360
gaagcaaatg	acgcgctatc	cggttcagacc	cagttcaacg	cggtgaacgt	actggtcaaaa	420
gcaggttctga	tgaattatca	ggttggtgcg	aatgacgggtg	aaactatcac	tatcgacctg	480
aagaaaaatc	attctgatac	ttctgggtctg	aatggtttta	acgttaaatg	taaaaggta	540
attaccaaca	aagctgcaac	ggtaagtgtc	ttactttctg	ctggcgcgaa	gttaaacacc	600
acgacaggct	ttatgatct	gaaaaccgaa	aatccctgtg	taactacgca	tgctgcattc	660
gataaaattag	ggaatggcga	taaaagtacc	gttggcgccg	tagattatac	ttacaacgct	720
aaatctgggtg	attttactac	accacaatct	actgctgtga	cggtgttaga	cgccgcggcg	780
caggctactg	attcagctaa	aaaacgtgat	gcgttagctg	ccacccttca	tgctgatgtg	840
ggtaaaattctg	ttaattgggtc	ttacaccaca	aaagatggta	ctgtttcttt	cgaaacggat	900
tcagcaggta	ataaccacat	cggtgggaag	caggcatacg	tagacgatgc	aggcaacttg	960
acgactaaca	acgtcggtag	cgcgactaaa	gctgatatga	aagcgtgct	taaaagccgcg	1020
gcaaatctgag	ctgcagctgc	ctctctgaca	ttcaatggca	ctgaataac	tatcgcaaaa	1080
gctacacgtga	gtaaaagatg	agttattgagc	cggttaaatcc	ctgggtggat	ttcttatcag	1140
attacacctta	attcccgctgt	actgagcaaa	actattgggt	ttaccgcggg	tgaatccagt	1200
gatgctgcga	agtccttatgt	ggatgataaa	gggtggtatta	ctaactgtgc	cgactataca	1260
gtctcttaca	gcgttaacaa	ggataacgcg	tcgtgactg	ttgcggggta	gactgtcgtc	1320
actgatacca	ataaagatta	tgctccagca	attggtaactg	ctgtaaatgt	gaactccgcg	1380
ggtaaaatca	ctactgagac	taccagtgt	gggttctgcga	cgacaaaccc	tatccagcagc	1440
ctggacgagc	ctatcagctc	catcgacaaa	ttccgttctt	ccctggggctg	tatccagaac	1500
cgctctggatt	ccgcagctcac	caacctgaac	aacaccacta	ccaaactctg	tgaagcagct	1560
tcocgtattc	aggacgcgga	ctatgcgacc	gaagtgtcca	acatgtcgaa	agcgagactg	1620
atccagcagg	ccggttaactc	cggtgctggca	aaagccaacc	aggtaccgca	gcaggtttctg	1680
tctctgctgc	agggttaa					1758

<210> 63  
 <211> 1758  
 <212> DNA  
 <213> Escherichia coli

```

<400> 63
atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaataa tatcaacaag 60
aaccagctct cgtctgcgag ttctatcgag cgtctgtctt ctggcttgccg tattaacagc 120
gcgaaggatg acgcgcgcgg tcaggcgatt gctaaccggt ttactttetaa cattaaagcg 180
ctgactcagg cggcccgtaa cgccaacgac ggtatttctg ttgcgcagac caccgaaggc 240
gcgctgtctg aatcaacaaa caacttacag cgtatccgtg aactgacggt tcaggccact 300
acaggggacta acccggtatc tgacctggag tgcattcgag tcaatccagg acgaaatcaa atctcgtctt 360
gatgaattcg acgcgcgtatc ggttgggtcg aatgacgctg aacacatcac gatcgacctg 480
gacggttcaa tgaataatca ggttgggtcg aatgacgctg aacacatcac gatcgacctg 480
aaaaaaatcg attctgatac tctgggtctg taaactctcg ctggcgcgaa gttaaacacc 600
attaccaaca aagctgcaac ggttaagtgt aataccttgt taactaccca tgctgcattc 660
acgacaggct tttatgatct gaaaaccgaa gttggcgccg tagattatac ttacaacgct 720
gataaattag ggaatggcga ctaaaatct cactaaatct cgggtgtaga cgcgcgcgct 780
aaatctgggt atttactac caaactgtat aaaaactgtat ccaccttca tgcgtcgctg 840
caggctgctg attcagcttc aaactgtgat ttaacccaca cggctgggagc caggctatcg tagacgatg aggcactctg 960
ggtaaatctg ttaattgggtc ttacacccaa cgtgatatag aagcgtctgt caaagcagcg 1020
tcagcaggta atatcaccat cgttgggaagc gcgactaaca acgctggtag cgcagctaaa tctcaatggca cagaatatac catcgcaaaa 1080
gcaactcctg cgacaaccac tccagtatgt agtatggagc actattgggt ttaccgcggg tgaatccagt 1140
gctacacgta gtaaaagtgt agtatggagc actattgggt ttaccgcggg tgaatccagt 1260
atttccttta attcgcggtg actgagcaaaa ggttgatata ctaactgtgc cgaactatca 1320
gatgctgcga agtctttatgt ggttgcacga attgtgactg ttgcgcggta tgcttcagcg 1380
gtctcttaca gcgttaacaa ggttgcacga attgtgactg ttgcgcggta tgcttcagcg 1380
actgtatgca ctactgagac taaccgtgct cactgcacaa caactgtgac ctaactgtgc cgaactgtgc 1440
ctgagcagcg caatcagctc caactgtgac ctaactgtgc cgaactgtgc 1500
cgtctggatt ccgcagctc ctaactgtgc ctaactgtgc cgaactgtgc 1560
tcccgatttc aggcgcgcga ctatgcgacc gaagtctcaa acatgcgcaa aggcagctatc 1680
attcagcagg ccggttaact cgtctgggca aaagctaac aggtaccgca gcaggttctg 1740
tctctgtgct aggggttaa
  
```

<210> 64  
 <211> 1758  
 <212> DNA  
 <213> Escherichia coli

```

<400> 64
atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaataa tatcaacaag 60
aaccagctct cgtctgcgag ttctatcgag cgtctgtctt ctggcttgccg tattaacagc 120
gcgaaggatg acgcgcgcgg tcaggcgatt gctaaccggt ttactttetaa cattaaagcg 180
ctgactcagg cggcccgtaa cgccaacgac ggtatttctg ttgcgcagac caccgaaggc 240
gcgctgtctg aatcaacaaa caacttacag cgtatccgtg aactgacggt tcaggccact 300
acaggggacta acccggtatc tgacctggag tgcattcgag tcaatccagg acgaaatcaa atctcgtctt 360
gatgaattcg acgcgcgtatc ggttgggtcg aatgacgctg aacacatcac gatcgacctg 480
gacggttcaa tgaataatca ggttgggtcg aatgacgctg aacacatcac gatcgacctg 480
aaaaaaatcg attctgatac tctgggtctg taaactctcg ctggcgcgaa gttaaacacc 600
attaccaaca aagctgcaac ggttaagtgt aataccttgt taactaccca tgctgcattc 660
acgacaggct tttatgatct gaaaaccgaa gttggcgccg tagattatac ttacaacgct 720
gataaattag ggaatggcga ctaaaatct cactaaatct cgggtgtaga cgcgcgcgct 780
aaatctgggt atttactac caaactgtat aaaaactgtat ccaccttca tgcgtcgctg 840
caggctgctg attcagcttc aaactgtgat ttaacccaca cggctgggagc caggctatcg tagacgatg aggcactctg 960
tcagcaggta atatcaccat cgttgggaagc gcgactaaca acgctggtag cgcagctaaa tctcaatggca cagaatatac catcgcaaaa 1020
acgcaagctg atgacagctg tctctgacaa ggttgatata aagcgtctgt ctaaatatac tatcgcacaa 1080
gcaactcctg cgacaaccct tccagtatgt agtatggagc actattgggt ttaccgcggg tgaatccagt 1140
ctgacagctg gtaaaagtgt agtatggagc actattgggt ttaccgcggg tgaatccagt 1260
attaccttca attcgcggtg actgagcaaaa ggttgatata ctaactgtgc cgaactatca 1320
gatgctgcga agtctttatgt ggtatgataa ggttgatata ttgcgcggta tgcttcagcg 1380
gtctcttaca gcgttaacaa ggtatgacgc cgtgtgactg ttgcgcggta tgcttcagcg 1380
  
```

actgataacca ataaagatta tgcctccagca attggtactg ctgtaaatgt gaactccgag 1440  
 ggtaaaaaata ctactcgagac taccagtgct ggtttctgcaa cgaccaaccc gcttgcgtgc 1500  
 ctggagcagcg ctatcagcttc catcgacaaa ttccgttctt cctctgggtgc tatccagaac 1560  
 cgctctggatt ccgacagtcac caacctgaac aacaccacta ccaacctgtc tgaagcgagc 1620  
 tcccgatttc aggcagccga ctatgcgacc gaagtgtcca acatgtcgaa agcgagatt 1680  
 atccagcagg ccggttaact cgtgctggca aaagccaaac aggtaccgca gcaggttctg 1740  
 tctctgtctgc agggtttaa

<210> 65  
 <211> 1758  
 <212> DNA  
 <213> Escherichia coli

<400> 65  
 atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaata tatcaacaag 60  
 aaccagctcg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120  
 gcgaaggatg acgcgcgggg tcaggcgatt gctaaccggt ttacttctaa cattaaagcg 180  
 ctgactcagg ctgcacgtaa cgccaaacgac ggtatttctg ttgcacagac cactgaagcg 240  
 cgctgtccgc aaatcaacaa caacttacag cgtatccgtg agctgacgtg tcaggcttct 300  
 accgggacta actctgattc ggaatctggac tccattcagg acgaaatcaa attccgcttc 360  
 gacgaaattg accgcgtatc cggctcagac cagttcaacg cggtgaacgt actggcaaaa 420  
 gacggttcga tgaaaattca ggttgggtcg aatgacggtg aaactatcac tatccagcgt 480  
 aagaataatg attctgatac tctgggtctg aatggtttta acgtaaatgg taaagtgatt 540  
 attaccaaca aagcgcgaac ggttaagtgat ttaactctta ctggcgcgca gtttaacacc 600  
 acgacaggtc ttatcgatct ggaagtcgaa aataacctgt taactaccga tgcgtattc 660  
 gataaat tag ggaatggcga taaagtcacc gttggcgcg tagattata ttacaacgct 720  
 aaatctgggt atttctaac caacaaatct actgctggtt cgggtgtaga cgcgcggcg 780  
 caggtactag attcagctaa aaaaogttag cgttagctg ccacccttca tgcgtatgtg 840  
 ggtcaactctg ttaattggtt ttacaccaca aaagatggta ctgtttctt cgaaaacggat 900  
 tcagcaggtt atataccatc cggtggaagc caggcatacg tagacgatg aggcaacttg 960  
 acgactcaaca acgctgttag cgcagctaaa gctgatatga aagcgtctg taaagccgag 1020  
 gcgaaggatg ctgcaggtgc ctctctgaca ttcaatggca ctgaataac tatcgcaaaa 1080  
 gcaactcctg cgacaaacctc tccagtactc ccgttaatcc ctggtgggat ttctttatcg 1140  
 gttacagtgta gtaaaagtgt agtatcgagc gaaaccaagc cggctgcgc gacatcttca 1200  
 attacccttta attccggtgt actgagcaaa actattgggt ttaccgcggg tgaactcag 1260  
 gatcgtcgca agcttctagt ggaatgataa ggtggtatta ctaacgttgc cgaatataca 1320  
 gctctttaca cgtttaacaa ggataaacgc tctgtgaact ttgcggggtt tgcctcagcg 1380  
 actgataaca ataaagatta tgcctcagca atttggtaact cgttaaatgt gaacrtccgc 1440  
 ggtaaaatca ctactgagac taacagtgct ggtttctgcaa cgaccaaccc gcttgcgtgc 1500  
 ctggagcagc ctatcagctc catcgacaaa tccggttctt cctctgggtgc tatccagaac 1560  
 cgctggagatt ccgcagctcac caacctgaac aacaccacta ccaacctgtc tgaagcgcag 1620  
 tcccgatttc aggcagccga ctatgcgacc gaagtgtcca acatgtcgaa agcgagatt 1680  
 atccagcagg ccggttaact cgtgctggca aaagccaaac aggtaccgca gcaggttctg 1740  
 tctctgtctgc agggtttaa

<210> 66  
 <211> 1758  
 <212> DNA  
 <213> Escherichia coli

<400> 66  
 atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaata tatcaacaag 60  
 aaccagctcg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120  
 gcgaaggatg acgcgcgggg tcaggcgatt gctaaccggt ttacttctaa cattaaagcg 180  
 ctgactcagg ctgcacgtaa cgccaaacgac ggtatttctg ttgcacagac cactgaagcg 240  
 cgctgtccgc aaatcaacaa caacttacag cgtatccgtg agctgacgtg tcaggcttct 300  
 accgggacta actctgattc ggaatctggac tccattcagg acgaaatcaa attccgcttc 360  
 gacgaaattg accgcgtatc cggctcagac cagttcaacg cggtgaacgt actggcaaaa 420  
 gacggttcga tgaaaattca ggtagggtcg aacgacggcc agactatcac tatgtatctg 480  
 aagaataatg actctgatac gctggggctg aatgggttta acgtgaatgg ttccggtatg 540  
 atagccaata aagcgcgac cttatgcgac ctgacagcag cgaataatgga tgcgtcaact 600  
 aataactata ctacaacaaa taatgcgctg actgcataca aggccttga tcaactgaaa 660  
 gatgttgaca ctgttactat caaagcagat gcagctcaaa ctgccacggt ctatacatc 720  
 aatgcactctg ctggttaact ctattcagtt aatgtatcga ataactctc agcaaaaagca 780  
 ggtgatgtag cagctagctt tctcccgccg cgtgggcgaa ctgctagtgg tgtttacaaa 840  
 gcgacagcag gtgaagtga cttttagtgt gatcgcaatg gtaaaattac atccggagga 900  
 caggaagcct atttaactag tgaatgtaac ttaactacaa acgatgctgg ttggtcgact 960

ggggctacgc ttgatgggttt attcaagaaa gctgggtgat gtcaatcaat cgggtttta 1020  
 aagactgcac cagtcacgat ggggggaaac actataaact ttaaaacggg tgcgtgatgt 1080  
 ggtgctgcga cgtgtaaacgc agggggtatcg ttcaactgata cagctagcaa agaaacccgt 1140  
 ttaaaataaag tgcctacagc taaacaaggc acagcagttg cagctaacgg tgatacatcc 1200  
 gcaacaatta cgtataaact tggcggttca acgtatcagg cgggtatttcg cgcaggtgac 1260  
 ggtactgcga gcgcaaaaata tgcggataat actgacggtt ctaatgcac agcaacatca 1320  
 acagatgctg actgggtaaat gactacaatt ggttcataca ccacgaagta ttcaatcgat 1380  
 gctacaacgc gcaaggttaac tgttgattct ggaactgggt cgggtaaaata tgcgcgcgaa 1440  
 gtccgggggtg aagtaatagt tagtgctaat ggtactttaa caacagatgc aactagcgaa 1500  
 ggcaacagta caaaagatcc actgaaagct ctggatgaag ctatcagctc ctatgcacaa 1560  
 ttcogttcaat cccctgggggc tatcaaaaac cgtttggatt ccgcgcgtcac caacctgaac 1620  
 aaacaccacta ccaactgtgc tgaagcgagc tcccgatttc aggaacgcga ctatggcacc 1680  
 gaagtgtcca caactgtgaa agcgagatgt atccagcagg ccggttaact cgtgctggca 1740  
 aaagccaaac aggtacgcga gcaggttctg tctctactgc aggggttaa 1788

<210> 67  
 <211> 1398  
 <212> DNA  
 <213> Escherichia coli

<400> 67  
 aacaaactct agtctctctct tagctctgct attgagcgct tctctcttgg tctgcgtatt 60  
 aacagcgcaa aagacgatgc agcagggtcag gcgattgcta accgttttca ggcgaatatt 120  
 aaaggtctga cccagggcttc ccgtaacgca aatgatgcta ttctgttgc gcagaccaact 180  
 gaaggtgcgc ttcttgaaat caacaataac ctgcagcgta ttctgtgaat ttctgtacag 240  
 gcaactaacg gtactaacct tgacagtgc ctgacctcca tccagtcga aactcacgac 300  
 cgtctgagtg aaattgatgc tttctctggt cagactcagt ttaacggcgt taaagtctgc 360  
 gctctctgac aggatgatgc tttctcaggtt ggtgcaaacg accgcgaaac aattactatt 420  
 aaactgcagg aaattaatct cgacacacgt ggattactgt gttttggatt taaagatcct 480  
 actcaaat taa aagcgcgaac ccgtgaaaca acctattttg gatcgacagt taagcttgc 540  
 cgcgttaata cgtctgatgc agatattaca gctacagtta aaggcactac gactccggcg 600  
 caacgtgacg cgaataattat gtcgtgatgt aactcgttaa agtgcgcgtt 660  
 tcagataaac cgtctgaaat tggttattat gaagtactgt tggaggatga tccgacatct 720  
 ctctgatcag taaagctgaa gctgggggct ctacgggtga cccagctcca agctgtgatt 780  
 ttaaaaggaag tcacaaacgt gaaagggaag ggggctattg atgttcagtt gggtagctag 840  
 accgcgaacg ctctatcac aggtgcaaaa ctctttaagt tagaagacgc caatgcgcaa 900  
 cgcgtcagct catcttgctt gatttggtgat gacggttaac agtatgcagc gaattgtgat 960  
 cagactgggt gaggcagtttc cgttaaaaaca atgtcttaca ctgatgcgtga cgtgtcaca 1020  
 cagcaaatg ttaaagttga actgggtgga agcgatggca aaacccgaag tgaactacga 1080  
 accgatgcaa aaacttacag tgttagtgat ttacaaggta agagcctgaa accgtattct 1140  
 attgcaagca ttcttcgca gaaaacagaa gatcctttgg ctgctatcga taaagcactg 1200  
 tctcaggttg aaactcttgc tteraaccta ggtgcataac aaatcggttt cgaactctgc 1260  
 atccacaac ctggcaaacac cgttaaacac ctgtctctct ccgtagccgc tatcgaaagt 1320  
 gctgactagc cgacggaagt gtctaacatg tctgtgcgc agatcctgca acaagcgggt 1380  
 acctctgttc tggcgcag

<210> 68  
 <211> 1479  
 <212> DNA  
 <213> Escherichia coli

<400> 68  
 aacaaactct agtctctctct gagctcgcgc attgaaacgt tctctcttgg cctgcgtatt 60  
 aacagtcgta aagatgacgc agcagggtcag gcgattgcta accgttttca agcaaatatt 120  
 aaaggtctga ctccaggttc ccgtaacgca aatgatgcta ttctgttgc gcagaccaact 180  
 gaaggtgcgc ttcttgaaat caacaataac ttacagcgta ttctgtgaat ttcgtagacg 240  
 gcaactaacg gtacaaactc tgactccgac ctgaattcaa ttcaggatga aattcacaa 300  
 cgtctgagtg aaattgatgc tgtttctaac cagacacaaat ttaattgggtt aaaagttctg 360  
 gctctgacg agactatgaa aattcaagta ggtgcgaacg atgggaaac cactgattct 420  
 gcccttgcaa aaattgatgc taaaacctgt gggcttgata accttagcgt agcaccagga 480  
 aaagttgcata tgtcctctgc ggttgcaact aagagcgaag ccgctcctga cttactaag 540  
 gtaaatgcga ctgatggtat tgtgggaagt gctaaagcat tccgttagcaa ttataaaaat 600  
 gtagtctgtg aaacttattt tggtaacggt aatgtacaag atacaaagga tacaactgat 660  
 gcgacggcta ctgcggaac aaaaagttat caagtacagg tggaaaggca gacttatttt 720  
 gtgtgtcaag ataataatc caacacgaac ggttttcat tattgaaca aaactctaca 780  
 ggttatgaaa aagttcaggt ggggtggtaag gatgttcagt tagcaaacct tgggtgtgct 840

[illegible]

WO 99/61458

- 1 -

PCT/AU99/00385

SEQUENCE LISTING PART

<110> THE UNIVERSITY OF SYDNEY

<120> ANTIGENS AND THEIR DETECTION

<130> REEVES

<140>

<141>

<160> 68

<170> PatentIn Ver. 2.0

<210> 1

<211> 1773

<212> DNA

<213> Escherichia coli

<400> 1

```
atgcgacgta tagaacgaat accgggggta tcggcgtaag cggggcaaaag tttaacgattt 60
attttttggc ttaatgacac gaacagcaac gaggaagggg agtatttcga ccgctagaaa 120
aaaaattctaa aggtttgtgag tgaccagacg ataacagggt tgacggcgac gaagccgaag 180
ggtggaagcc caatacttaa accgtagact tgaaaacagg aaaaatgaatc atggcacaag 240
tcattaatat caacagcctc tcgctgatca ctcaaaaataa tatcaacaag aaccagtctg 300
cgctgtcgac ttctatcgag cgctctctct ctggtctgcy cattaacagc gctaaagatg 360
acgctgcggg ccaagcgatt gctaaccgct tcacttctaa catcaaaagg ctgactcagg 420
ccgcacgtaa cgccaacgac ggtatttctc tggcgcgac cactgaaggc gcactgtctg 480
aaatcaacaa caacttcgag cgtgttctgt aactgaccgt tcaggccact accggtacta 540
actctgattc tgacctgtct tcaatacagg acgaaatcaa atcccgctct gatgaaattg 600
accgcgtatc cggtcagact cagttcaacg cgtttaatgt tctttccaaa gatggttcaa 660
tgaaaattca gggttggtcg aatgatggtc aaactatctc catcgatctg aagaaaattg 720
attcttcaac ttgggggctg aatggcttct cagtttctaa aaactctctt aatgtcagca 780
atgctatcac atctatccg caagccgcta gcaatgaacc tgttgatggt aacttcggtg 840
atactgatga gtctgcagca atcgagcca aattgggggt ttccgatacg tcaagcctgt 900
cgctgcacaa catccttgat aaagatggta aggcacacgc tgattatggt gttcagtcag 960
gtaagacttt ctatgctgct tctgttaatg ccgcttcagg taaagtaacc ttaaacacca 1020
ttgatgttac ttatgatgat tatcggaacg gtgttgacga tgccaagcaa acaggtcagc 1080
tgatcaaaat ttacagcagat aaagacggcg cagctcaagg ttttgtcaca cttaaggcca 1140
aaaactattc tgctggtgat gcggcagaca ttcttaagaa tggagcaaca gctcttaagt 1200
taactgatct gaatttaagt gatgttactg atactaatgg taaggtaacc acaactgcga 1260
ctgagcaatt tgaaggtgct tcaactgagg atccgctggc gcttctggat aaagctattg 1320
catcagtcga caaattccgg tcttctctag tgcccgatga gaacctctc gattccgcta 1380
tcaccaacct gaacaaaccc accaccaacc tgtctgaagc gcagtcocgt attcaggacg 1440
ccgactatgc gaccgaagtg tccaacatgt cgaagcgca gatcatccag caggcaggta 1500
actccggtgt gtctaaagcg aaccagggtac cgcagcaagt tctgtcactg ttacaaggct 1560
aatggcctta acctgcctga ccccgccacc ggccgggggtt tttctgtccg caatttaccg 1620
```

ataaccccc aataaccctt catttcaccc actaatcgtc cgattaaaaa cccctgcagaa 1680  
 acggataatc atgccgataa tccatataac gcaggcgctgt ttatcggtgaa ttcactctat 1740  
 accgctgaag gtgtaatgga taaacactcg ctg 1773

<210> 2

<211> 500

<212> DNA

<213> *Escherichia coli*

<400> 2

aacagcctct cgtggtatc tcagaacaac atcaacaaaa accagtcttc aatgtctact 60  
 gccattgagc gtctgtcttc cgggtctggt atcaacagcg caaaagatga cgtcgtctggc 120  
 caggcgattg ccaaccgctt cacccttaac atcaaaggtc tgactcaggc agctcgtaac 180  
 gccaacgacg gtatctccgt tgcacagacc actgaaggcg cactgtctga aatcaacaac 240  
 aacctgcagc gtatccgtga gctgactgtt cagtcttcta cgggtactaa ctctgaaatcc 300  
 gatctgaact caatccagga cgaataataa tcccgctctg acgaaattga ccgcgtatcc 360  
 ggtcagacc agttcaacgg cgtgaacgtg ctggcaaaag acggctccat gaaaattcag 420  
 gttggcgca acgatgtgga aaccatcacc atcgacctga aaaaaattga ctctcttact 480  
 ttaaacctga ctgggtttaa 500

<210> 3

<211> 500

<212> DNA

<213> *Escherichia coli*

<400> 3

ctcagtatgc tgtcacccgc agtacagggt ccgtaactta cgtaccagat acagatccgt 60  
 ccgcgactgg tgatatgttt tctgcttatg ttgatgatgc aggtacattg acaactgatg 120  
 caaacaataa tgtaaaaat tttgcccaca ctaattgtag cgtcacgaac gacagtgggt 180  
 cagctattta cgcactgaa gcgggcaaat tgactactga agcgtctaca gctgctgaaa 240  
 ctaccgctaa cccactgaaa gccctggacg atgcaatcag ccagatcgac aaattccgtt 300  
 cttctctggg tgctgtacag aaccgtctgg attctcggtt aaccaacctg aacaacacca 360  
 ccaccaacct gtcggaagcg cagtcccgta ttcaggacgc cgactatgcg accggaagtgt 420  
 caaatatgtc taaagccgag atcatccagc aggcgggtaa ctccgtgttg gctaaagcta 480  
 accaggttcc tcagcaggtt 500

<210> 4

<211> 399

<212> DNA

<213> *Escherichia coli*

<400> 4

agcctgtcgc tgttgacca gaataacctg aacaaatctc agtcttctct gagctccgcc 60  
 attgagcgct tctctctctg cctgcgtatt aacagtgcga aagatgacgc agcaggtcag 120  
 gcgattgcta accgttttac agcaaatatt aaaggctctga ctcaggcttc ccgtaacgcg 180  
 aatgatggta tttctgttgc gcagaccact gaaggcgcgc tgaatgaaat taacaacaac 240  
 ctgcagcggt tacgtgaact gactgttcag gcaactaacg gtactaactc tgacagcgat 300  
 ctttcttcta tccaggtgta aattactcaa cgtctggaa gaaattgacg tgatctctgag 360



caaaactcagt ttaacggcgt gaaagtcctt gctgaaaat

399

<210> 5

<211> 417

<212> DNA

<213> Escherichia coli

<400> 5

gcacgttagt tggttaacggt gcaacttacg atgttagtgc agatggtaaa acgataacgg 60  
agactgcttc tggttaacaaat aaagtcattg atctgagcaa atcagaagggt ggtagccgca 120  
ttcttggtaaa cgaagatgca gcaaaatcgt tgcaatctac caccaaccgg ctcgaaacta 180  
tcgacaaagc attggctaaa gttgacaatc tgcgtcttga cctcgggtgca gtacaaaacc 240  
gtttcgactc tgcctatcac aaccttggca acaccgtaaa caacctgtct tctgcccgta 300  
gccgtatcga agatgctgac tacgcgacgg aagtgctctaa catgtctcgt gcgcagatcc 360  
tgcaacaagc gggtacctct gttctggcgc aggttaacca gaccacgcag aacgtac 417

<210> 6

<211> 950

<212> DNA

<213> Escherichia coli

<400> 6

aacaaaaacc agtctgcgct gtcgacttct atcgagcgcc tctctcttgg tctgcgtatt 60  
aacagcgcta aagatgacgc cgcggggccag gcgattgcta accgctttac ttctaaccatc 120  
aaaggtctga ctcaggccgc acgtaacgcc aacgacgcta ttctcttggc gcagacgggt 180  
gaaggcgcgc tgcagagatg taacaacaac ttgcagcgta ttctggaact gaccgttcag 240  
gcctctaccg gcacgaactc tgattccgac ctgtcttcta ttcaggacga aatcaaatcc 300  
cgtcttgatg aaattgaccg tgtatctggt cagaccgatg tcaacgggtg gaacgtgctg 360  
tcgaaaaacg attcgatgaa gattcagatt ggtgccaatg ataaccagac gatcagcatt 420  
ggcttgcaac aaatcgacag taccactttg aatctgaaag gatttaccgt gtccggcatg 480  
gcggatttca gcgcggcgaa actgacggct gctgatggta cagcaattgc tgcctgggat 540  
gtcaaggatg ctggggggttaa acaagtcaat ttactgtctt acactgacac cgcgtctaac 600  
agtactaaat atgcggtcgt tgattctgca accggtaaat acatggaagc cactgtagtc 660  
attaccggta cggcggcggc ggtaactgtt ggtgcagcgg aagtggcggg agccgctaca 720  
gccgatccgt taaaagcact ggatgccgca atcgctaaag tcgacaaatt ccgctctctc 780  
ctcggtgccg ttcaaaaaccg tctggattct gcggtcacca acctgaacaa caccaccacc 840  
aacctgtctg aagcgcgact ccgtattcag gacgcgact atgcgacga agtgctcaac 900  
atgtcgaaag cgcagattat ccagcaggcg ggcaactccg tgcgtctctaa 950

<210> 7

<211> 1212

<212> DNA

<213> Escherichia coli

<400> 7

aacaaaaacc agtctgcgct gtcgacttct atcgagcgcc tctctcttgg tctgcgtatt 60  
aacagcgcta aagatgacgc cgcggggccag gcgattgcta accgctttac ttctaaccatc 120  
aaaggtctga ctcaggccgc acgtaacgcc aacgacgcta tctctcttgg gcagaccact 180

gaagcgcgcg tgtctgaaat caacaacaac ttgcagcggt tgcgtgagtt gaccgttcag 240  
 gcgcagcagc ggactaactc tgattctgac ctgtcttcta ttcaggacga aatcaaatcc 300  
 cgtctggatg aaattgatcg cgtttccggt cagaccaggt tcaacggcgt gaattgtgctg 360  
 gcgaaagatg gttcgtatga gattcaggtt ggccggaatg atgggcagac tatcagcatt 420  
 gatttgcaga agattgactc ttctacatta ggactgaacg gtttctccgt ttccgggtcag 480  
 tcaactaacg ttagtgtatc cactactcaa attaccgggt ccgcggggac aaaacctggt 540  
 ggtgttgatt tcactgctgt tgcgaaagat ctgactactg cgacaggtaa aacagtgcag 600  
 gtttctagcc tgacgttaca caacactctg gatcggaagc ggggtgctac atcacagttc 660  
 tgcgttcaat ccggcaatga ttctactcc gcgtcgatta atcacacaga cggcaaatgc 720  
 acgttgaata aagccgatgt cgaatacaca gacaccgata atggactaac gactcgcggt 780  
 actcagaagc atcaactgat taaagttgcc gctgactctg acggctcggc tgcgggatac 840  
 gtaacattcc aaggtaaaaa ctacgctaca acggtttcaa cggcacttga tgataactac 900  
 gcggcaaaa ccaacagataa taaagtgttt gttgaattat caacagcaaa accgactgca 960  
 cagttctcag gggctttctc tgctgatcca ctggcacttt tagacaaagc tattgcacag 1020  
 gttgatactt tccgctcttc cctcgggtgcg gtgcgaaacc gtcgtgattc cgcagtaacc 1080  
 aaacctgaaca acaccacac caacctgtct gaagcgcatg ccggtattca ttcgctccgac 1140  
 tatgtacacg aagtgtccaa catgtcgaaa gcgcagatca tccagcaggc aggtaactcg 1200  
 gtgtgtgcca aa 1212

&lt;210&gt; 8

&lt;211&gt; 1647

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 8

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaata tatcaacaag 60  
 aaccagtcgt cgtctgcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120  
 gcgaaggatg acgcgcggcg tcaggcgatt gctaaccggt ttacttctaa cattaaggcg 180  
 ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcgcagac caccgacgcg 240  
 gcgctgtccg aaattaacaa caacttacag cgtattctgt aactgacggt tcaggctctc 300  
 accgggacta actctgattc ggactctggc tccattcagg acgaaatcaa atcccgctc 360  
 gacgaaattg acccgctatc cggtcagacc cagttcaacg gcgtgaaagt actggcaaaa 420  
 gacgggttcga tgaataatca ggttggtgcg aatgacggcc agactatcac tattgatctg 480  
 aagaaaaatt actctgatac gctggggcgt aatgggttta atgtgaacgg caaaggggaa 540  
 acggcttaata cggcgcaac cctgaaagat atgtctggat tcacagctgc ggcgacacca 600  
 gggggaaactg ttggtgtaac tcaatatac gacaaatcgg ctgtagcaag tagcgtagat 660  
 attcataagt ctgttgcgtg cgcagatgga aataaagtta caactagcgc cgatgttgtt 720  
 tttggtacac caccgcgtgc tgtaacctat acctacaata aagacactaa ttcataatcc 780  
 gccgcttctg atgatatttc cagcgctaac ctgggtgctt tcttcaatcc tcaggccgga 840  
 gatacgacta aagctacagt tacaattggt ggcaagatc aagatgtaaa catcgataaa 900  
 tccggttaatt taactgtgc tgatgatggc gcagtacttt atatggatgc taccggtaac 960  
 ttaactaaaa ataagtctgg ttggtgataca caagctactt ttgctaaact tgctactgct 1020  
 actggtgcta aagccgcgac catccaaact gataaaggaa cattcacag tgacggtaca 1080  
 gcggttgatg gtgcataat gtccattgat accaatatc ttgcaaatgc agtaaaaaat 1140  
 gacactata ctgcactgt aggtgtcaag acttatagc taacaacagc ttctgctgct 1200  
 gcagacacgc ctatatagc caatgggggt ctcagtgata ctccgcaac ttactatgca 1260  
 caagctgatg gaagtatcac aactactgag gatcgcggtc ccggtaaaact ggtctacaaa 1320  
 ggttccgatg gtaagttaac aacggatagc actagcaaac cagaataaac atcagatccg 1380

ctggcagctc ttgacgacgc tatcagccag atcgacaaat tccgctctcc cctgggtgcg 1440  
gtgcaaaaacc gtctcgattc cgcagtgacc aacctgaaca acaccactac caacctgtct 1500  
gaagcgcagt cccgtattca ggacgcgcac tatgcgaccg aagtgctcaa catgtcgaaa 1560  
gcgcagatta tccagcagcg cggttaactcc gtgctggcaa aagctaacca ggttccgcag 1620  
caggttctgt ctctgtcgca ggggttaa 1647

&lt;210&gt; 9

&lt;211&gt; 1758

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 9

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaata tatcaacaag 60  
aaccagctcg cgctcgctag ttctatcgag cgctcgtctt ctggcttgcg tattaacagc 120  
gcgaaggatg acgcccgggg tcaggcgatt gctaaccgtt ttacttctaa cattaaaggc 180  
ctgactcagg ctgcacgtaa gcaccaacgac ggtatttctg ttgcacagac caccggaaggc 240  
gcgctgctcg aaatacaaaa caacttacag cgtatccgtg agctgacggt tcaggcttct 300  
accggaacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgctct 360  
gatgaatttg accgcgtatc cggccagacc cagttcaacg gcgtgaacgt actggcaaaa 420  
gacggttcga tgaaaaattca ggttggtgcg aatgacggtg aaactatcac tatcgacctg 480  
aagaaaatcg attctgatac tctgggtctg aatgggttta acgtaaatgg taaaggtaact 540  
attaccaaca aagctgcaac ggtaagtgat ttaacttctg ctggcgcgaa gttaaacacc 600  
acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgcattc 660  
gataaattg ggaatggcga taaagtcacc gttggcgcg tagattatac ttacaacgct 720  
aaactcgttg attttactac caccaaatct actcgtggta cgggtgtaga cgcccgggcg 780  
caggctactg attcagctaa aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840  
ggtaaatctg ttaatgggtc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900  
tcagcaggta atatcccat cggtggaagc caggcatatc tagacgatgc aggcaacttg 960  
acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgctgct taaagccgct 1020  
agcgaaggta gtgacgggtc ttctctgaca ttcaatggca ctgaatatac tatcgcaaaa 1080  
gcaactcctg cgacaacctc tccagtagct ccgttaatcc ctggtgggat tacttatcag 1140  
gtacagtgta gtaaatgagt agtattgagc gaaccacaaag cggctgcgcg gacatcttca 1200  
attaccttta attccgggtg actgagcaaa actattgggt ttaccggcgg tgaatccagt 1260  
gatgctgcga agtcttatgt ggtatgataa ggtggtatta ctacgctgc cgactataca 1320  
gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgcggggta tgcttcaggc 1380  
actgatacca ataaagatta tgctccagca attggtactg ctgtaaatgt gaactccgcg 1440  
ggtaaaatca ctactgagac taccagtgtc ggttctgcaa cgaccaaccc gcttgctgcc 1500  
ctggacgacg ctatcagctc catcgacaaa ttccgttctt cctctgggtg tatccagaac 1560  
cgtctggatt ccgcagctac caacctgaac aacaccacta ccaacctgtc tgaagcgcag 1620  
tcccgtattc aggacgcgca ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatt 1680  
atccagcagg ccggtaaact cgtgctggca aaagccaaac aggtaccgca gcaggttctg 1740  
tctctgctgc aggggttaa 1758

&lt;210&gt; 10

&lt;211&gt; 1383

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 10

aacaaatctc agtctctctc tagctctgct attgagcgct tctctcttgg tctgcgtatt 60  
 aacagcgcaa aagacgatgc agcaggctcag gcgattgcta accgttttac ggcaaatatt 120  
 aaaggtctga cccaggcttc cgttaacgca aatgatggta ttctcgttgc gcagaccact 180  
 gaaggtgcgc tgaatgaaat taacaacaac ctgcagcgta ttctgtgaact ttctgttcag 240  
 gcaactaacg gtactaactc tgacagcgat ctctcttcta tccaggctga aattactcaa 300  
 cgtctggaag aaattgacgc gtatatctgag caaactcagt ttaacggcgt gaaagtcctt 360  
 gctgaaaaa atgaaatgaa aattcagggtt ggtgctaagt atgggtgaac catcactatc 420  
 aatctggcaa aaattgatgc gaaaactctc ggccctggacg gtttaatat cgtatggcgcg 480  
 cagaaagcaa caggcagtgta cctgatttct aaatttaaag cgacaggtac tgataattat 540  
 gatgttggcg gtaaaactta taccgtgaat gtggagagcg gcgcggttaa gaatgatgct 600  
 aataaagatg tttttgtaag cgcagctgat ggatcgctga cgacagtag tgataactaa 660  
 gtatccggtg aaagtattga tgcaacagaa ctacgcaaac ttgcaataaa attagctgac 720  
 aaaggtctca ttgaatacaa gggcattaca ttactaaca acactggcgc agagcttgat 780  
 gctaatggta aaggtgtttt gaccgcaaat attgatggtc aagatgttca attactatt 840  
 gacagtaag caccacgggg tgccggcgca acaataacta cagacacagc tgtttacaaa 900  
 aacagtcgag gccagttcac cactacaaaa gtggaaaaa aagccgcaac actctctgat 960  
 ctggatctta atgcagccaa gaaaacagggt agcactttag ttgtaaatgg cgccacctac 1020  
 aatgtcagc cagatggtaa aacggttaact gatactactc ctgggtgccc taaagtgatg 1080  
 tatctgagca aatcagaagg tggtagcccg attctggtaa acgaagatgc agcaaaatcg 1140  
 ttgcaatcta ccaccaaccc gctcgaaaact atcgacaagg cattggctaa agttgacaat 1200  
 ctgcgtctcg acctcggtgc agtacaacac cggttcgact ctgccatcac caaccttggc 1260  
 aacaccgtaa acaactctgc ttctgccctg agccgtatcg aagatgctga ctacgcgacc 1320  
 gaagtgtcta acatgtctcg tgcgcagatc ctgcaacaag cgggtacctc tgttctggcg 1380  
 cag

1383

&lt;210&gt; 11

&lt;211&gt; 2013

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 11

atggcacaag tcattaatc caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60  
 aaccagctcg cgctctcgag ttctatcgag cgtctgtctt ctgggttgcg tattaacagc 120  
 gcgaaggatg acgcgcgggg tcaggcgatt gctaacgcgt ttacttctca cattaaaggc 180  
 ctgactcagg ctgcacgtaa cgccaacgac ggtatttcg ttgcacagac cactgaaggc 240  
 gcgctgtccg aaattaacaa caacttacag cgtattctgt aactgacggt tcagggttct 300  
 accgggaact actccgattc ggaatcgagc tccattcagg acgaaatcaa attcccgctg 360  
 gacgaaattg accgcgtatc cggccagacc cagttcaacg gcgtgaacgt gctgtccaaa 420  
 gatggctcga tgaaaaattca ggtcgccgcg aacgatggcg aaacgattac ttatgatctg 480  
 agaaaaattg actctgatac gctgaattcg gctgggttta acgttaacgg taaaggttct 540  
 gtacgcaata cagctcgagc aagcgacgat ttaaaacttg ctgggttccac taagggcacc 600  
 acagatacca atggcgtagc cgcgtataca aacacaatta gtaatgacaa agccaaagct 660  
 tccgactctg tagctaatat caccgatgga tcagtgatca ctgggggagg ggcaaacgct 720  
 ttggcgctgg ctgcaaaaga tggttacacc tatgatgcag caagtaaatc ttatagttt 780  
 gcctgcagat gtgcgcatc agcgaagacg ttaagcatca ttaatccaaa caccgggtgat 840  
 tctgcgagc cgacagtgac tattggtggt aaagagcaga aagttaatat tggccaggat 900  
 ggaaaaatta ctgcgcgaga tgataatcgc acgctgtatt tagataaaca tgccaaactg 960

acaaaaacga atgcaggtaa cgataccgca gcgacttggg atggtttaat ttccaacagc 1020  
 gattctaccg gtgcgggttcc agttgggggtt gcaactacaa ttacaattac ttctgggtaca 1080  
 gcttccggaa tgtctgttca gtccgcagga gcaggaattc agacctcaac aaattctcag 1140  
 attcttcgag gtggtgcatt tgcggctaaag gtaagtattg agggagggcg tgctacagac 1200  
 attttggtag caagtaattg aaacataaca gcggctgatg ttagtgcatt ttactttgat 1260  
 ggcactactg ttggatttcc tacaacggct ggaggaaata cagctgtctc gttagataat 1320  
 ttaattgcta acagtaagga tgctacctta acgtaacct caggtaccgg ccagacaact 1380  
 gtttatagca caacaggaag tggcgctcag ttcaccagtt tagcaaaagt agacacagtc 1440  
 aatgtcacca acgcacatgt cagtgcgcga ggtatggcaa atctgacaaa aagcaatttt 1500  
 accattgata tgggcggtag aggtacagta acttacacag ttccaatgg ggatgtgaaa 1560  
 gtgctgcgaa atgtctgatg ttatgtcgaa gatggtgcac ttccagccaa tgctacaaaa 1620  
 gatgtaacct actttgaaca aaaaaatggg gctattacca acagcacggg tggtagccatc 1680  
 tatgaacacg ctgatggtaa gttaacaaca gaagctacta ctgcatccag ttccaccgcc 1740  
 gatccccga aagctctgga cgaagccatc agctccatcg acaaatccg ctccctccctc 1800  
 ggtgcgggtg aaaaaccgtct ggattccgcg gtcaccaacc tgaacaacac cactaccaac 1860  
 ctgtccgaag cgcagtcctc tattcaggag gccgactatg cgaccgaagt gtccaacatg 1920  
 tcgaaagcgc agatcatcca gcaggccggt aactccgtgc tggcaaaagc taaccaggta 1980  
 ccgcagcagg ttctgtctct gctgcagggt taa 2013

&lt;210&gt; 12

&lt;211&gt; 1263

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 12

atggcacaa gtcattaatac caacagcctc tcgctgatca ctcaaaaataa tatcaacaag 60  
 aaccagtcgt cgtgtctcag ttctatcgag cgtctgtctt ctggcttctg tattaacacg 120  
 gcgaaggatg aaaccccggt tcaggcgatt gctaaacgct ttactttcaa cattaaaggc 180  
 ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcgcagac caccgaaggc 240  
 cgctgtctcg aaattacaac caacttacag cgtgtgcgtg agctgactgt tcaggcgacc 300  
 accggtacta actctgagtc tgacctgtct tctatccagg acgaaatcaa attctgcctg 360  
 gaagagattg atcgtgtttc aagtcagact caatttaacg gcgtgaatgt ttggctaaa 420  
 gatgggaaaa tgaacattca ggttggggca aatgatggac agactatcac tattgatctg 480  
 aaaaagatgt attcatctac actaaacctc tccagtttgg atgtacaaa ctggggcacc 540  
 agtgttaag atggggccac catcaataag caagtggcag taggtgtctg cgactttaaa 600  
 gataaagctt caggatcggt aggtacccta aaattagttg agaaagacgg taagtactat 660  
 gtaaatgaca ctaaaagtag taagtactac gatgccgaag tagatactag taagggtaaa 720  
 attacttcaa actctacaaa tgaagtgga actactccta ctgcagcgac ggaagtaact 780  
 actgttggcc gcgatgtaaa attgtagtct tctgcactta aagccaacca atcgctgtctc 840  
 gtgtataaag ataaaagcgg caatgatgct tatatcattc agaccaaga tgtaacaact 900  
 aatcaatcaa ctttcaatgc cgctaatac agtgaatgct gtgttttctc tattggtgca 960  
 tctacaaccg cgccaagcaa ttaaacagct aaccgcgcta aggtctctga tgatgcaatt 1020  
 gcatctgttg ataaattccg ctcttctctc ggtgcggtt agaaccgtct ggattctgcc 1080  
 attgccaaac tgaacaacac cactaccaac ctgtctgaag cgcagtcctg tattcaggac 1140  
 gctgactatg cgaccgaagt gtccaacatg tcgaaagcgc agattatcca gcaggccggt 1200  
 aactcgtgct tggcaaaagc caaccaggta ccgcagcagg ttctgtctct gctgcagggt 1260  
 taa 1263

<210> 13  
 <211> 1368  
 <212> DNA  
 <213> Escherichia coli

<400> 13  
 aacaaatctc agtctctctc gagctccgcc attgaacgtc tctctctctg cctgcgtatt 60  
 aacagtgtca aagatgacgc agcagggtcag gcgattgtcta accggtttac agcaaatatt 120  
 aaaggctctga ctcaggcttc ccgtaacgcg aatgatgtta ttctgtgtgc gcagaccact 180  
 gaagggtgcgc tgaatgaaat taacaacaac ctgcagcgtg tacgtgaact gactgttcag 240  
 gcaactaacg gtactaactc tgacagcgat cttctctcta tccaggctga aattactcaa 300  
 cgtctggaag aaattgaccg tgtatctgag caaactcagt ttaacggcgt gaaagtctct 360  
 gctgaaaata atgaaatgaa aattcaggtt ggtgctaag atggtgaaac catcactatc 420  
 aatctggcaa aaattgatgc gaaaactctc ggcctggcag gttttaatat cgtatggcgcg 480  
 cagaaagcaa ctggcagtgga cctgatttct aaatttaaac gcacaggtac tgataactat 540  
 gatgtgtggc gtgatgtcta tactgttaac gtatagatagc gagctgttaa agatactaca 600  
 gggaatgata tttttgttag tgcagcagat ggttcactga caactaaatc tgacacaaac 660  
 atagctggta cagggtatga tgcacagca ctgcagcag cggctaagaa taagcacag 720  
 aatgataaat tcacgttttaa tggagtgtga ttcaacaaca caactgcagc ggtatggcaat 780  
 gggaatgggt tatattctgc agaaatgtat ggtaagttag tgacatttac tgtacagat 840  
 gctgacaaa aagcttcttt gattacgagt gagacagttt aaaaaatag cgtgtggcctt 900  
 tatacgacaa ccaaaagtga taacaaggct gccacacttt ccgactctga tctcaatgca 960  
 gctaagaaaa caggaagcac gttagtgttt aacggtgcaa cttacagatg tagtgacag 1020  
 ggtaaaacga taacggagac tgcttctggt aacaataaag tcatgtatct gagcaaatca 1080  
 gaaggtggta gcccgattct ggtaaacgaa gatgcagcaa aatcgttgca atctaccacc 1140  
 aacccgctcg aaactatcga caaagcattg gctaaagttg acaatctgcg ttctgacctc 1200  
 ggtgcagtac aaaaccgttt cgactctgct atcaaccaac ttggcaacac cgtataacaac 1260  
 ctgtctctcg cccgtagccg tatcgaagat gctgactacg gaccggaagt gcttaacatg 1320  
 tctcgtgcgc agatcctgca acaagcgggt acctctgttc tggcgacg 1368

<210> 14  
 <211> 1788  
 <212> DNA  
 <213> Escherichia coli

<400> 14  
 atggcacaag tcattaatac caacagcttc tcgctgatca ctcaaaaata tatcaacaag 60  
 aaccagctcg cgtcgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120  
 gcgaaggatg acgcagcggg tcaggcgatt gctaaccggt tcaactctaa cattaaggcg 180  
 ctgactcagc cggcccgtaa cgccaacgac ggtatctccg ttgcgcagac caccgaagcg 240  
 cgcgtctccg aaatcaacaa caacttacag cgtatccgtg aactgacggt tcaagctctc 300  
 accgggacta actccgattc ggtatctggac tccattcagg acgaaatcaa atcccgctcg 360  
 gacgaaattg accgcgtatc tggccagacc cagttcaacg cgtggaacgt actggcgaaa 420  
 gacggttcaa tgaaaattca ggttggtcgc aatgacggcc agactatcac gattgatctg 480  
 aagaaaattg actcagatac gctggggctg aatgggttta acgtgaatgg ttccggtacg 540  
 atagccaata aagcggcgac cattagcgac ctgacagcag cgaaaatgga tgcgtcaact 600  
 aatactataa ctacaaacaaa taatgcgctg actgcataaa aggcgcttga tcaactgaaa 660  
 gatgggtgaca ctgttactat caaagcagat gctgctcaaa ctgccacggt ttatacatc 720

```

aatgcatcag ctggttaactt ctcatcagtt aatgtatcga ataatacttc agcaaaaagca 780
gggtgatgtag cagctagcct tctcccgccg gctgggcaaa ctgctagtgg tgtttataaa 840
gcagcaagcg ctgaagttaa ctttgatgtt gatgcgaatg gtaaaatcac aatcggagga 900
cagaaagcat atttaactag tgatggtaac ttaactacaa acgatgctgg tggtcgcact 960
cgggctacgc ttgatggttt attcaagaaa gctgggtgatg gtcaatcaat cgggttttaa 1020
aagactgcac cagtcacgat ggggggaaca actataact ttaaaacggg tgcctgatgt 1080
gatgctgcga ctgctaaccg aggggtatcg ttcactgata cagctagcaa agaaaccgtt 1140
ttaaataaag tgcctacagc taacaaggcg aaagcagttg cagctgacgg tgatacatcc 1200
gcaacaatta cctataaato tggcgttcag acgtatcagg ctgtatttgc cgcagggtgac 1260
ggtagtgcta gcgcaaaaata tgccgataaa gctgacgttt ctaatgcaac agcaacatc 1320
actgatgctg atgggtgaaat gactacaatt gggtcatata ccacgaagta ttcaatcgat 1380
tgtaacaacg gcaaggtaac tgttgattct ggaactggta cgggttaaata tgcgcggaaa 1440
gtaggggctg aagtataatg tagtgctaatt ggtagcttaa caacagatgc aactagcgaa 1500
ggcacagtaa caaaagatcc actgaaagct ctggatgaag ctatcagctc catcgacaaa 1560
ttccgtttctt cctcgggtgc tatccagaac cgtctggatt ccgcagtcac caacctgaac 1620
aacaccacta caaacctgct cgaagcgagc tcccgatcc aggcagcgca ctatgcgac 1680
gaagtgtcca acatgtcgaa agcgcagatc attcagcagg ccggtaactc cgtgctggca 1740
aaagccaacc aggtaccgca gcaggttctg tctctgctgc aggggttaa 1788

```

&lt;210&gt; 15

&lt;211&gt; 1653

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 15

```

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaata tatcaacaag 60
aaccagtcct cgtctgcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
cggaaggatg acgcccaggg tcaggcgatt gctaacggtt ttacttctaa cattaaggcg 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttcgg ttgcgcagac catgaagggt 240
gcgctgtccg aaatcaacaa caacttcagc cgtatttcgt agctgacggt tcaggctctt 300
accgggacta actccgattc tgacctggac tccatccagg acgaaatcag gctctgctgt 360
gacgaaattg acccgctatc cggctcagac cagttcaacg cgtgtaacgt gctcgcgaaa 420
gacgggttca tgaataatca ggttggtgag aatgacggcc agactatcac gattgatctg 480
aagaaaaatt actcagatgc gctggggctg agtgggttta atgtgaatgg tggcggggct 540
gttgctaaac ctgctgcac taaagctgac ttggtagctg ctaatgcaac tgggttaggc 600
aacaatata ctgtgagtgc gggttacgat gctgctaaag cgtctgattt gctggctgga 660
gttagtgatg gtgatactgt tcaggcaacc attaaataac gcttcgggac ggcgctagt 720
gcaacgaatt acaagtatga cagtgcgaag aagctctact cttttgatac caacaaggct 780
tcagctgcgg atgttcagaa atatttgacc cggggcggtt gtgataccgc taaggcgact 840
attactatcg atggtctcgc acaggatggt cagatcagca gtgatggtaa aattacgtca 900
agcaatggag ataaacttta cattgatata actgggcgct taacgaaaaa cggctttagt 960
gcttctttga ctgaggctag tctgtccaca cttgcagcca ataataccaa agcgacaacc 1020
attgacattg gcggtaacct tatctccttt accggttaata gtactacgac gaacactatt 1080
actttacag taacagggtgc aaaagttgat caggcagctt tcgataaagc tgtatcaacc 1140
tctggaaaag atgttgattt cactaccgca ggttatagcg tcgacggcgc aagtcgctgt 1200
gtacaaaaag gtgtgtctcc ggtttatatt gataacaacg gggcggtgac cacatctgat 1260
actgtagatt tttatctaca ggatgatggt tcagtgacta accggcagcg taaggcagtt 1320
tataaagatg ctgacggtaa attgacgaca gatgctgaaa ctaagctgac aaccaccgac 1380

```

```

gatccctga aagctctgga cgaagccatc agctccatcg acaaattccg ctcctccctc 1440
ggtgcggtgc agaaccgtct ggattccgcg gtcaccaacc tgaacaacac cactaccaac 1500
ctgtctgaag cgcagtcctcg tattcaggac gctgactatg cgaccgaagt attccaaatg 1560
tcgaaagcgc agatcatcca gcaggccggg aactccgtgc tggcaaaagc taaccaggta 1620
ccacagcagg ttctgtctct gctgcagggt taa 1653

```

&lt;210&gt; 16

&lt;211&gt; 1689

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 16

```

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaata tatcaacaag 60
aaccagtcct cgctgtctcg ttctatcgag cgtctgtctt ctggcttgcc tattaacagc 120
gcgaaggatg acgcgcgagg tcaggcgatt gctaaccgtt ttacttctaa cattaaggcg 180
ctgactcagg ctgcacgtaa cgcacaacgac ggtattttctg ttgcacagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtgtgcgtg aactgacctg tcaggcaacc 300
accggtacca actccagtc tgacctggac tctatccagg acgaaaattaa atcccgctcg 360
gacgaattg atcgctgtat cggtcagacc cagttcaacg gcgtgaacgt gctggcaaaa 420
gacggttcca tgaaaattca ggttggcgcg aacgatggcc agaccatcac tatcgacctg 480
aagaagattg actcttctac cttgaacctg acaggtttta acgttaacgg ttctggttct 540
gtggcgaaata ctgcagcaac taaagctgat ttaaccgctg ctcaactctc tgcaccgggt 600
gcagcagacg caaatggtag agttacttat actgtcagtg ctggttataa agaattccact 660
gctgcagatg ttattgctag catcaaaagc ggcagtgtct cgactcttgc aattacttga 720
accatttaata atggcttcgg tgattccagt gcgctgactt ccaatgacta tacttatgac 780
ccagcaaaag gcgacttcac ttacgacgta gcttcaagcg ccaataatac tctgcccagc 840
gttcagtcct tcttgacgcc gaaagcagggt gataccgcaa atctgaaagt aaccgttggt 900
acgacatcgg ttgatgtcgt tctggccagt gatggttaaga ttacagcaaa agatggttct 960
gcattatata tcgacagtac aggtaacctg actcagaaca gtgctggctt gacctctgct 1020
aaactggcta ctctgactgg ccttcagggc tctggtgttg ctccaacct cactactgaa 1080
gatggcacta atattgatat tgctgttaac ggtaatatg gtctgaccgg tgttcgtatc 1140
agtctgatt ctctgcagtc agcgactaaa tctacgggct ttactgttgg tactggcgct 1200
acaggtctga cctgaggtag tgatggtaaa gtgactatcg gcgggactac tgctcagctc 1260
tacacagca aagatggttc cctgactact gataaaccca ctaaaactgta tctgcagaaa 1320
gatggctctg taaccaacgg ttacaggtaaa gcggtctatg tagaagcgga tggatattc 1380
actaccgacg ctgcaaccaa agccgcaacc accaccgacg cgctgaaagc cctggatgag 1440
tcgaactagcc agatcgataa gttccgttca tccctgggtg ctatccagaa ccgtctggat 1500
tcccggttca ccaacctgaa caacaccact accaacctgt ctgaagcgga gtcccgatt 1560
caggacgccc actatgcgac cgaagtgtcc aacatgtcga aagcgcagat cattcagcag 1620
gccggttaact ccgtgctggc aaaagccaac caggtagccg aacaggttct gtctctgctg 1680
cagggtctaa 1689

```

&lt;210&gt; 17

&lt;211&gt; 915

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 17



gcgctgtcga cttctatcga ggcctctctt tctggtctgc gtattaacag cgcataaagat 60  
gacgctgcgg gccaggcgat tgctaaccgc ttactttcta acatcaaaagg tctgactcag 120  
gcgcgacgta acgccaacga cggatattct ctggcgcgaga cggctgaagg cgcgctgtca 180  
gagattaaca acaacttgca gcgtattcgt gaactgaccg ttcaggcctc taccggcagc 240  
aactctgatt ccgacctgtc ttctattcag gacgaaatca aatcccgctt tgatgaaatt 300  
gaccgtgtat cctgtcgagc ccaggtcaac ggtgtgaacg tgctgtcgaa aaacgattcg 360  
atgaagattc agattgggtc caatgataac cagacgatca gcattgggctt gcaacaatec 420  
gacagttacca ctttgaatct gaaaggattt accgtgtccg gcattggcgga ttctagcgcg 480  
gcgaaactga cggctgctga tggtagacga attgctgctg cggatgtcaa ggatgctggg 540  
ggtaaacaa gtcattttact gtcttacct gacaccgctg ctaacagtac taaatcgcg 600  
gtcgttgatt ctgcaaccgg taaatacatg gcagccactg tagtcattac cagtacggcg 660  
gcggcggtaa ctgttgggtc aacggaagtg gcgggagcgg ctacagccga accgttaaaa 720  
gcactggatg ccgcaatcgc taaagtcgac aaattccgct cctccctcgg tgcggttcaa 780  
aacctgtcgg attctggcgt caccacactg aacaacacca ccaaccaact gtctgaagcg 840  
cagtcgccga ttcaggacgc cgactatgcg accgaagtgt ccaacatgct gaaagcgcg 900  
attatccagc aggcg 915

&lt;210&gt; 18

&lt;211&gt; 1665

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 18

atggcacaa gtcattaatc caacagcctc tcgctgatca ctcaaaaata tatcaacaag 60  
aaccagctcg cgcctgcgag ttctatcgag cgtctgtctt ctggccttgc tattaaacag 120  
gcgaagtgat acgcgcgagg tcaaggcgatt gctaaccggt ttacttctaa tattaaggcg 180  
ctgactcagg ctgcacgtaa cgccaatgac ggtattttct tgcacagac cactgaaggc 240  
gcgctgtccg aaatcaacaa caacttacag cgtattctgt aactgacggt tcaggccaact 300  
acagggacta actccgattc tgacctggac tccatccagg acgaaatcaa actctcgctg 360  
gacgaaattg accgcgtatc cggtcagacc cagttcaacg cgcgtgaact gctgtccaaa 420  
gatggttcaa tgaaaaattc ggtcggcgca aatgatgggt aaaccatcac gattgatctg 480  
aagaaaaattg actctgatac gctgaatctg gctggtttta acgtgaatgg cgaagggtga 540  
acagccaata ctgctgcaac acttaaagat atggttgggt taaaactcga taatacgggg 600  
gtcactacag ctggagttaa tagatatatt gctgacaaag ccgtgcgaag tagcaccgat 660  
attttgaatg cggtagctgg tgttgatggc agtaaaagtt ccacggagggc agatgttgg 720  
tttgggtgag ctgcccctgg taacgccagt gaatatact atcataaaga tactaacaca 780  
tatacggctt ctgcttcagt tgatgcgact caactggcgg cattcctgaa tcttgagcg 840  
ggtggtacca ctgctgcaac agtaagtatt ggcaacggta caacagctca agagcaaaaa 900  
gtcattattg ctaaagatgg ttctttaact gctgctgat acggtgccgc tctctactct 960  
gatgatactg gtaacttaag taaaactaac gcaggcactg atactcaagc taaactgtct 1020  
gacttaatgg caaacaatgc taatgccaaa acagtcatta caacagataa aggtacattt 1080  
actgataata cgacaaagtt tgatggggta gatatttctg ttgatgtctc acggtttgct 1140  
aacgcggtta aaatcgagac ttacactgca actgttgggt taactttacc tgcacatat 1200  
acagtcata atggcactgc tgcacagcgt tattttagtc atggaaaaag gagcaaaact 1260  
cctgcgagat attttgctca agctgatggc actattacta gtggtaaaaa tgcggctacc 1320  
agtaaaagcta tctatgtaa tgccaatggg aacttaacga ctaatacaac tagtgaatct 1380  
gaagctacta ccaaccgcgt ggcagcattg gatgacgcta tcgcgtctat cgacaaattc 1440  
cgttcttccc tgggtgctat ccagaaccgt ctggattccg cagtcacaaa cctgaacaa 1500

accactacca acctgtctga agcgcagtc cgtattcagg acgccgacta tgcgaccgaa 1560  
 gtgtccaaaca tgtcgaaagc gcagatcatt cagcaggccg gtaactccgt gctggcaaaa 1620  
 gccaccagg taccgcagca ggttctgtct ctgctgcagg gttaa 1665

&lt;210&gt; 19

&lt;211&gt; 1842

&lt;212&gt; DNA

<213> *Escherichia coli*

&lt;400&gt; 19

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaata tatcaacaag 60  
 aaccagctcg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120  
 gcgaaggatg acgcccaggg tcaggcgatt gtaaccggtt ttactcttaa cattaaaggc 180  
 ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcgcagac cactgaaggc 240  
 gcgctgtccg aaattaaaca caacttacag cgtattcgtg aactgacgtt tcaggcgacg 300  
 accggaacta actccacctc tgacctggac tccatccagg acgaaatcaa atcccgctct 360  
 gacgaattg accgcgtatc tggtcagacc cagttcaacg cgctgaacgt gctgtctaaa 420  
 gatggctcga tgaaaattca ggtcggcgcg aacgatggcg aacgattac tattgatctg 480  
 aagaaaaatt actctgatac gctgaatctg gctggtttta acgttaacgg taaaggttct 540  
 gttagcgata cgcgtcgac tacagataat ctgacattgg ctggttttac agcgggtact 600  
 aaagctcgtg atggcaccgt aacttatagc aaaaatgtcc agtttgccgc cgcgactgca 660  
 agcaatgtac tggctctgc taaagatggc gacgaatta cgctcgctgg taataacggc 720  
 acaggtatag ctgcaactgg ggggacttat acttatcata aggactctaa ctcatcacg 780  
 tttagcgcaa cggctgcac taaagattct ctgttgagca cactggcacc aaacgctggc 840  
 gatacattta ccgctaaagt gactattggt tctaaatcgc aagaagttaa cgttagcaaa 900  
 gatgtaacga ttacatccag cgatggtaag cgctgtatt tagatgagaa gggcaacctg 960  
 acccaaacag gtatggcac aaccaaagct gcaacctggg ataactgat ggcataatac 1020  
 gatactacag gcaaaagatc ctatggtaac tctgcggcag cagctgttgg gacagtaac 1080  
 gaagcaaaag gaatgaccat caattctgct ggtggtaatg ctacaggtgtt aaaagacgcg 1140  
 gcttataatg cgcgatatgc gacctcaatt actactggta ctccgggtga tgcgggagcc 1200  
 gcggggagcc gtgcaactgc gggtaatgcc gcggtgggag cgctgggccc aacggcagtt 1260  
 gataatacca cggcagatgt tgccgatata tctatctcag ctctcgaaat ggcgagcatt 1320  
 cttcaggata aagatttcac cttaagtgtg ggtagtata ctacaacgt caccagcaat 1380  
 gctgtcacta tcaatggcaa agcagcaaac attgatgaca cggcgcaat gacagacaa 1440  
 accagtaaa ttgtcaatta tttcgtcat actaacggta gcgtgactaa cgatacaggc 1500  
 tccactattt atgcgacaga agatggtagc ctgaccaccg atgcagcaac caaagcgaa 1560  
 accaccgccg atccctgaa agctctggac gaagccatca gctccatcga caaattccgc 1620  
 tctctccctg gtgcgggtga aaacgctctg gattccggg tcaccaacct gaacaacacc 1680  
 accaccaacc tgtctgaagc gcagtccegt attcaggacg cgcactatgc gaccgaagt 1740  
 tccaacatgt cgaaagcgca gattatccag caggccggta actccgtgct ggcaaaagct 1800  
 aaccaggatc cacagcaggt tctgtctctg ctgcagggtt aa 1842

&lt;210&gt; 20

&lt;211&gt; 1731

&lt;212&gt; DNA

<213> *Escherichia coli*

&lt;400&gt; 20

atggcacaag tcattaatac caacagcctc togetgatca ctcaaaaataa tatcaacaag 60  
 aaccagctct cgcgtctgag ttctatcgag cgtctgtctt ctggcttgcy tattaacagc 120  
 gcgaaggatg acgcgcgagg tcaggcgatt gctaacctgt ttacttctaa cattaaagcg 180  
 ctgactcagc cggcccgtaa gcccaacgac ggtatttctg ttgcygcagc caccgcaagc 240  
 gcgctgtccg aaattacaac caacttacag cgtgtgcgtg agctgactgt ctaggcgacc 300  
 accggtacca actccagctc tgatctggac tctatccagg acgaaatcaa atcccgctctg 360  
 gacgaaattg accgcgtatc cggtcagacc cagttcaacg gcgtgaacgt gctggcaaaa 420  
 gacgggtcca tgaaaaatca ggttgccgcy aatgatggcc agaccatcac tatcgacctc 480  
 aagaagattg actcttctac gttgaaactg actgggttta acgtgaatgg ttctggttct 540  
 gtggcggaata ctggcgcgac taaagcggat ttggctgctg ctgcaattgg taccctctgg 600  
 cgagcagatt ctacaggtgc cattgcttac acagtaagtg ctgggctgac taaaactaca 660  
 gccgcagatg tactgtctag cctcgctgat ggtacgacta ttacagccac aggcgtgaaa 720  
 aatggctttg ctgcaggagc cacttccaat gctataaac ttaacaaaga taataatata 780  
 ttacttatg acacgactgc tacgacagct gagctgcagt cttacctgac tccgaaagcg 840  
 ggccgacact caacattcag tgttgaaatt ggtggtaact cacaagactg ctgctgtcc 900  
 agtgatggca aactcactgc taaggatggc tctaagcttt acattgatac aactggtaat 960  
 ttaactcaga atggttgtaa taacggtgtt ggaacactcg cgggaagcgac tctgagtgtt 1020  
 tttagctctg acaaaaaatg tttaacggct gttaaatcca caattactac agctgataac 1080  
 acttcgattg tactgaatgg ttcaagcgat ggtactggta agtctggtag tgaaggtagc 1140  
 attgctgtta caggcgctgt aattagttca gctgctctgc aatctgcaag caaaacgact 1200  
 ggtttcactg ttggtacagt agacacagct ggttatatct ctgtaggtag tgatgggagt 1260  
 gttcaggcat atgatctgcg gacttctggc aacaaagctt cttacaccaa cactgacggt 1320  
 acactgacta ctgataacac cactaaactg tatctgcaga aagatggctc tgaataacaa 1380  
 ggttcaggta aagcgttcta tgtagaagcg gatggtgatt tcaactaccg cgtgcaacc 1440  
 aaagccgcga ccacaccgca tccgctggcc gctctggatg acgcaatcac ccagatcgac 1500  
 aagttccgtt catccttggg tgctatccag aaccgtctgg attctgcagt caccaacctg 1560  
 aacaacacca ccaccaacct gtctgaagcg cagtcgccga ttcaggacgc cgactatgcy 1620  
 accgaagtgt ccaatatgtc gaaagcgag atcctccagc aggcgggtaa cctcgctctg 1680  
 gcaaaagcca accaggtacc gcagcaggtt ctgtctctgc tgcagggtta a 1731

&lt;210&gt; 21

&lt;211&gt; 1380

&lt;212&gt; DNA

<213> *Escherichia coli*

&lt;400&gt; 21

aacaaatctc agtcttctct gagtccgcgc attgaacgct tcttctcttg cctgcgtatt 60  
 aacagtgcta aagatgacgc agcaggtcag gcgatttgcta accgttttac agcaaatatt 120  
 aaaggctctg ctacaggttc ccgtaacgcy aatgatggta ttctgttgc gcagaccact 180  
 gaagggtgcg tgaatgaaat taacaacacg ctgcagcgta ttctgtaact ttctgttcag 240  
 gcaactaacg gtactaactc tgacagcgat ctttcttcta tccaggtcta aattactcaa 300  
 cgtctggaag aaattgaccg tgtatctgag caaactcagt ttaacggcgt gaaagfcctt 360  
 gctgaaaaata atgaaatgaa aattcaggtt ggtgctaagt atggtgaaac catcactatc 420  
 aatctggcaa aaattgatgc gaaaactctc ggcttgagcy gttttaatat cgatggcgcy 480  
 cagaagacaa ccggcagtgta cctgatttct aaatttaagc cgacaggtac tgataattat 540  
 caaattaacg gtactgataa ctatactgtt aatgtagata ttggcgtagt acaggataaa 600  
 gatggcaaac agtgttttgt gagtactgcy gatgggttcc ttacgaccag cagtgtactc 660  
 caattcaaga ttgatgcaac taagcttgca tgggtgcta aagatttagc tcaagggaaat 720

aagattgtct acgaaggat cgaatttaca aataccggca ctgtcgctat agatgccaa 780  
 ggtaatggta aattaaccgc caatgttgat ggtaaggctg ttgaattcac tatttcgggg 840  
 ggtactgata catcagggtac tagtgcaacc gttgccctta cgacagccct atacaaaaat 900  
 agtgccgggc aattgactgc acaaaaagtt gaaaaataag cagcgacact atctgatctt 960  
 gatctgaacg ctgccaaaga aacagggaagc acgttagttg ttaacggtgc aacttacgat 1020  
 gttagtgca atggtaaaac gataacggag actgctctct gtaacaataa agtcctgtat 1080  
 ctgagcaaat cagaagggtg tagccccatt ctggtaaacg aagatgcagc aaaaatcgtt 1140  
 caatctacca ccaaccgcct cgaaactatc gacaaagcat tggctaaagt tgacaactct 1200  
 cgttctgacc cctgtgcagt acaaaaacgt ttcgactctg ccatcaccac ccttggcaac 1260  
 accgtaaca accgtgtctt tgcctctagc cgtatcgaag atgctgacta cgcgaccgaa 1320  
 gtgtctataa tgtctctgct gcagatcctg caacaagcgg gtacctctgt tctggcacag 1380

&lt;210&gt; 22

&lt;211&gt; 1767

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 22

atggcacaag tcattaatac caacagcctc tcgtgatca ctcaaaaata tatcaacaag 60  
 aaccagtcgt cgtctgcag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120  
 gcgaaggatg acgcagcggg tcaggcgatt gctaaccgtt ttacttctaa cattaaaggg 180  
 ctgactcagg cggcagctaa cgccaacgac ggtatctctc tggcgacagc caccgaaagg 240  
 gcgctgtctg aaatacaaca caacttacag cgtgtacgtg aactgacctg tcaggcaacc 300  
 accggtacta actccgactc cgacctggct tctattcagg acgaatataa atcccgctgt 360  
 gatgaaattg accgctgctc tggtcagact cagttcaacg cgtggaacgt gctggcaaaa 420  
 gacggttcca tgaaaaatca ggtagggtgt aacgacggcc agactatcac tattgacctg 480  
 aaaaaaatcg actctgatac tctgggctgt aatggtttta acgtgaatgg ttctgggacg 540  
 attaccaaca aagcagcaac tgtcagtgat gttactcgcg caggcgggat attggtgaat 600  
 ggtgcctatg atataaaaac cactaacaca cgcgtgacta caactgatgc ctctcgcaaa 660  
 ttgaatgatg gtgatgtgtg tactatcaat aatggtaagg atactgccta taaataataa 720  
 gctgctacag tggggtttac gacggatgtc tccatctccg gggatcctac gctgctgacg 780  
 gctactgcta ataaactgc cgtgatgca ctttcggcgt ctttactatg ctgcggcggg 840  
 aaaactgtta atggtctctg gactacgaat gatggtacgg taaaatttga taccgatgcc 900  
 gatggtaaag tttctatttg tgggtgtgct gcttatgtag atgcagcagg caacctgacc 960  
 actaacgcag caggtatgac gactcaagca atttggttac tctgtctgca 1020  
 tctgtactcg gtaagggtgg atccctgacc tttggtgaca cgacgtataa aattggtcag 1080  
 ggtacggctg ggggtgatcc tgatgacgct tcagatgatg tactggggac cattttctac 1140  
 tctaaatcag taagcaagga tgtgtgtctt gctgatacta agcaactagg taacacgca 1200  
 acagttgatt tcaactccgg tatcatgact tcaaagggtta gtttcgatgc aggtacatca 1260  
 actgatacat tcaaaagtgc agatggtgct atcacaaaaa ctaaagaata caccactctt 1320  
 tatgtctgaa ataaagatac tgggtgaagt accgttgcgt attatgtctg ggtagatagc 1380  
 gccgataagg ctgttgatga tactaaatata aaaccgacta tcggcgcgac agttaacctg 1440  
 aatctctcag tgaatttgac cactgatacc accaggtcag gcacagcaac caaagatcct 1500  
 ctggctgcc cggacgtctg tatcagctcc atcgacaaat tccgttctac cctgggtgct 1560  
 atccagaacc gtctggatcc cgcagtcacc aaactgaaca acaccactac caactgtctc 1620  
 gaagcgcagt cccgtattca ggacgcgacg tatcgacgag aagtgtccaa catgtcgaaa 1680  
 gcgcagatta tccagcaggc cggtaactcc gtgctggcaa aagccaacca ggtaccgcag 1740  
 cagggttctgt cctctgataa ggggttaa 1767

- 15 -

<210> 23  
 <211> 1383  
 <212> DNA  
 <213> Escherichia coli

<400> 23  
 aacaaaaacc agtctgcgct gtcgacttct atcgagcgcc ttctctctgg tctgcgtatt 60  
 aacagcgcta aagatgacgc tgcgggcccag gcgattgcta accgcttcac ttctaacatc 120  
 aaagggtctga ctcaggccgc acgtaacgcc aacgacggta ttctctctggc gcagaccact 180  
 gaaggcgcgct gttctgagat taacaacaac ttgcagcgctg tgcgtgagtt gactgtacag 240  
 gcgacgaccg ggaactaact tgattctgac ctgtcttcta tccaggatga aatcaaatcc 300  
 cgtttaacgc aaattgaccg tgtatctggt cagactcagt ttaacggcgt gaacgtactg 360  
 gctaagaatg acaccctgtc tattcaggta ggtgcaaatg acggtcagac tatcaatatt 420  
 gacctgcagc aaatcgattc tcatacactg ggtctggatg gtttcagcgt taaaaataat 480  
 gatgcagtga aaaccagtg cgcgtgaat actcttggg ggggggcagg tctgtgttgc 540  
 gtcgacttcg caacaaccag tttagctgct atcactggtc tcggtagcgg tgctatcagc 600  
 gaaattgcta aagacgataa tgggtattac tacgcgcagt tcacagggac tacgggtaat 660  
 actgctgatg gttactatgc tgctgatata gacaaggcta ccggtgaggt cgctctgaaa 720  
 gatggtaacg tagatacacc gacaggtacg ccaacgacga caagcacata tgacttcaac 780  
 gacgctgtgc aaaccgttcc ctttggcact gatgctgcaa cagccggtat cagactgtgt 840  
 gcttctctct ttaacttcoa ggatgagaaa ggcaatgata ctgctactta tgcaatcaaa 900  
 gcacaagatg gcagcctgta tgcgcgaac gttgatgagg ctaccggtaa agtcactgtc 960  
 aaaaacgcca gctatactga tgctgacggc aaagcagtgta ccgatgcgcc gtgaaaaactg 1020  
 ggtgggtgaca atggcaaac cgaaattgtt gtgcagtctg cgtcaggtaa aacttaacgat 1080  
 gctggtgcac tgcaaaacgt tgatctctcc agtgcaccca acacggtaac cgcaatcccg 1140  
 aacggtaaaa ccacgtctcc gctggctgcc ctgacgacg caatcagcca gatcgacaaa 1200  
 ttccgctect cctcgggtgc ggtgcagaac cgtctggatt ccgcggtcgc caactggaac 1260  
 aacaccacta ccaacctgtc tgaagcgag tccgtattc aggacgctga ctatgcgacc 1320  
 gaagtatcca acatgtcgaa agcgcagatc atccagcagg caggtaactc cgtgctgtcc 1380  
 aaa 1383

<210> 24  
 <211> 1197  
 <212> DNA  
 <213> Escherichia coli

<400> 24  
 gcgctgtcga cttctatcga gcgcctctct tctgggtctgc gcattaacag cgctaagat 60  
 gacgctgcgg gccaaagcat tgctaaccgc ttcacttcta acatcaaaag tctgactcag 120  
 gccgcacgta acgccaacga cgggtatttct ctggcgacga ccaactgaagg cgcactgtct 180  
 gaaatcaaca acaacttgca cgggtgttctg gaactgacgc ttcaggccac taccggtact 240  
 aactctgatt ctgacctgtc ttcaatacag gacgaaatca aatcccgctc cgtatgaatt 300  
 gaccgggtat cgggtcagac tcagttcaac ggcggtaatg ttctttccaa agatgggtca 360  
 atgaaaaattc aggttggtgc gaatgatggt caaactatct ccacgatctc gaagaaaatt 420  
 gattcttcaa ctttggggct gaatggcttc tcagtttcta aaaactctct taatgtcagc 480  
 aatgctatca catctatccc gcaagccgct agcaatgaac ctggtgatgt taactcgggt 540  
 gatactgatg agtctgcagc aatcgacgcc aaatggggg ttcccgatac gtcaagccgtg 600

tcgctgcaca acatccttga taaagatggt aagccaacag ctgattatgt tgttcagtea 660  
 ggtaaagact tctatgtctg tctctgtaat gccgcttcag gtaaagtaac cttaaacacc 720  
 attgatgtta cttatgatga ttatgcgaac ggtgttgacg atgccaagca aacaggctcag 780  
 ctgatcaaag tttagcagaga taaagacggc gcagctcaag gttttgtcac acttcaaggc 840  
 aaaaactatt ctgctggtga tgcggcagac attcttaaga atggagcaac agctcttaag 900  
 ttaactgac tgaaatttaag tgatgttact gataactatg gtaaggtaac cacaactcgc 960  
 actgagcaat ttgaaggctg tccaactgag gatccgctgg cgtctctgga taaagctatt 1020  
 gcatcagctg acaaatccg gtctctctca ggtgcgctgc agaaccgtct cgattccgct 1080  
 atccaacac tgaacaacac caccaccaac ctgtctgaag cgcagtcctc ttacaggcact 1140  
 gccgactatg cgaccgaagt gtccaacatg tcgaaagcgc agatcatcca gcaggca 1197

&lt;210&gt; 25

&lt;211&gt; 1674

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 25

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60  
 aaccagctct cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120  
 gcgaaggatg acgcccaggg tcaggcgatt gctaacctgt ttacttctaa cattaaaggc 180  
 ctgactcagg ctgcacgtaa cgccaacgac ggtattctct ttgcacagac cactgaaggc 240  
 gcgctgtcgc aaatcaacaa caactacag cgtattctgt aactgacgtg tcaggccaact 300  
 acagggacta actccgattc tgacctggac tccatccagg acgaaatcaa atctcgtctg 360  
 gacgaaattg accgcgtatc tggtcagacc cagttcaacg gcgtgaacgt gctgtctaaa 420  
 gatggctcga tgaanaattca ggtcggcgcg aacgatggcg aaacgattac tattgatctg 480  
 aagaaaattg actctgatac gctaaatctg gctgttttta acgtgaattg tgctggctct 540  
 gttgataatg ccaaggcgac tggcaaatgat ctactgatg ctggttttac ggcaaggcga 600  
 gctgatgcta atggcaaaat cacttatacc aaagacacgc ttactaaatt cgacaaggcg 660  
 acagcggctg atgtattggg caaagcggct gctggcgata gcattacctg tcggggcact 720  
 gatactggct taggagctgc tgcgtatgcc tcgacttaca cctacaatgc agccaataag 780  
 tcttacactt ttgatgctac tgggtgttgc aaggcggatg ctggaacggc actgaaaggg 840  
 taacttaggc cactcaacac cggtaaaatt aatctcggtg tcaccgagca agaagctaac 900  
 attgccaaag atggctccat caccgatacc aatggcgatg cgtgtgatct cgatagtacc 960  
 ggcaacttaa ccaaaaatac cgcgaaattg ggggtctgct ataaagcaac tgtagataaa 1020  
 ctgtttctgt gtgctcagga tgcaacgata acctcgata gggcactgac agctaaattc 1080  
 gatcaaaactg ctggtaacctg tgatttcaaa ggcgcgtcta tttctgtcta tgcaatggca 1140  
 tcaaccttaa ataatggttc ctatacagcc aacgtagggt gtaaggctta tgccgttaac 1200  
 gctggcgagc ttacagacagc tggcgagatg gtgtataaag ataccactgg cgactgacg 1260  
 actgaagatg acgaaacgct taccgcgacc tactacggtt ttgctgatgg taaagtttct 1320  
 gacggtgaag gttctactgt ctataaagct gctgatggtt ccactactaa agatgcgact 1380  
 accaagtctg aagcaaccac tgacctctgt aaagcccttg acgacgcatc cagccagcatc 1440  
 gacaattcc gctctctccct cgttgcgctt caaaacgctc tggattccgc cgtcaccac 1500  
 ctgaacaaca ccaactacaa cctgtctgaa gcgcagtcct gtattcagga cggcgactat 1560  
 gcgaccgaag tctccaacat gtcgaaagcg cagatcattc agcaggccgg taactccgtg 1620  
 ctggcaaaaag caaccaggt accgcagcag gttctgtctc tgcgtcaggg ttaa 1674

&lt;210&gt; 26

&lt;211&gt; 1365

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 26

aacaaatctc agtcttctct tagctctgct attgagcgtc tctcttctgg cctgcgtatt 60  
aacagtgccta aagatgacgc agcaggtcag gcgattgcta accgttttac gccaaatatt 120  
aaagggtcga ctcagggttc ccgtaacgag aatgatggta tttctgttgc gcagactact 180  
gaaggtgcgc tgaatgaaat taacaacaac ctgcagcgty tactgtaaact gactgttcag 240  
gcaactaacg gtactaactc tgacagcgat ctttcttcta ttcaggcaga aattactcaa 300  
cgctcggaag aaattgacgc tgtatctgag caaactcagt ttaacggcgt gaaagtcctt 360  
gccgaaaata atgaaatgaa aattcagggtt ggtgctaagt atgggggaaa catcactatc 420  
aatctggcaa aaattgatgc gaaaactctc ggcttggaag gctttaatat cgatggcgcg 480  
cagaaagcaa ctggcgagta cctgatttct aaatttaaa gcacaggtac tgataattat 540  
caaatatacg gtactgataa ctatactggt aatgtagata gtggagcagt tcaaatagag 600  
gatggtgacg caatttttgt tagcgctacc gatggttctc tgactactaa gagtgatata 660  
aaagtgcgtg gtacaggtat tgatgcgact gggcttgcaa aagcgcgagt tctcttagct 720  
aaagatgcct caattaataa ccaagggtatt actttcacca acaaaaggac tgatgcattt 780  
gatggcagtg gtaacggcac tctaacgcgt aatattgatg gcaaaagatgt aacctttact 840  
attgatgcga cagggaagga cgcaacatta aaaacgtctg atcctgttta caaaaatagt 900  
gcaggctcagt tcactacaac taagggtgaa aacaagccg ctacagcacc ggatctggac 960  
ttaataaacg ctaaaaaagt gggtagttct ttagttgtaa atggcgctga ttaataagtt 1020  
agcgctgatg gtaagacagt aactgggctt ggcaaaacta tgbatctgag caaatcagaa 1080  
gggtggtagc cgattctggt aaaagaagat gcagcaaaaa cgttgcaatc tactaccaac 1140  
ccgctcgaaa ccatacgaaa ggcattggct aaagttgaca atctgcgttc tgacctcggt 1200  
gcagtacaaa accgttttga ctctgtctac accaaccttg gcaacaccgt aaacaacctg 1260  
tctctgcgcc tgagccgcat cgaagatgct gactacgcga cgaagtgctc taacatgtct 1320  
cgtgcgcaga tctgcacaa agcgggtacc tctgttctgg cgcag 1365

&lt;210&gt; 27

&lt;211&gt; 1740

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 27

atggcacaaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60  
aaccagtcgt cgctgtcgag ttctatcgag cgtctgtctt ctggcttggc tattaacagc 120  
cgcaaggatg acgcccagag tcaggcgatt gctaaccggt ttacttctaa cattaaggcg 180  
ctgactcagg ctgcacgatg cgccaacgat ggtatttctg ttgcacagac cactgaaggc 240  
gcgctgtccg aaatcaacaa caacttacag cgtatccgtg aactgacggt tcaggcttct 300  
accgggactg actccgattc ggaatctggac tccattcagg acgaaatcaa atcccgctctg 360  
gacgaaattg accgcgtatc ttggccagacc cagttcaacg cgtgaaacgt actggcgaaa 420  
gacggttcaa tgaaaattca ggttgggtcg aatgacggcc agactatcad gattgatctg 480  
aagaaaaattg actctgatac gctggggctg agtgggttta atgtgaatgg tagcggggct 540  
gtggctaata ctgcagcgac taaatctgat ttggcagcag ctcaactctt ggctccaggt 600  
actcgtgatg ctaattggtac agttacctat actgttggcg caggcctgaa aacatctaca 660  
gctgcagatg taattgcgag tttggctaat aacgcaaaag ttaatgccac aattgcaaat 720  
ggttttggat cgccaacagc tacagattat acatacaaca gcgctacagg cgattttaca 780  
tatagtgcga ctattgcagc tgggtacaaat tctgggtgata gtaacagtg ctagttacaa 840

```

tccttctcga caccaaaagc gggcgatact gctaacttaa acgttaaaat tgggtctacg 900
tcaattgacg ttgtattggc tagcgacggg aaaattaccg cgaagatagg ttacagaacta 960
ttatttgacg tagatggtaa cctcactcaa acaaatgctg ggaactgtcaa agcagccact 1020
cttgatgcac tgaactaaaa ctggcataca acaggcacac cgagtggcgt atctacggta 1080
attacaactg aagatgaaac aacctctact ctggctggcg gtactgatgc tactactttc 1140
ggtgcaatca ctgtagcaaa tgcaagaatg agtgctgagt ctcttcaatc ggcaactaag 1200
tccacaggat tcacagttag tgttggagct actggtacca gcgcaggcga tattaaggtt 1260
gatagtaaa gtatagtaca acaacacaca ggtacaggtt ttgaagacgc tcaccacaa 1320
gctgattggt cactgactac cgataatata accaatctgt ttttgcaaaa agacgggaact 1380
gtgaccaatg gttcaggtaa agcagttctat gtttcagcgg atgggtaattt tactactgac 1440
gctgaaacta aagctgcaac caccggcgat ccaactgaaag ctctggacga agcgatcagc 1500
tccatcgaca aattccgttc ttccctcggt gcggtgcгаа accgtctgga ttccgcagtc 1560
accaacttga caaacaccac tactaacctg tctgaagcgc agtcctcgat tcaggacgct 1620
gactatgcga cgaagtgttc caatatgtcg aaagcgcaga tcatccagca ggccggttaac 1680
tccgtgctgg caaaagctaa ccaggtaccg cagcaggttc tgtctctgct gcagggttaa 1740

```

&lt;210&gt; 28

&lt;211&gt; 1233

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 28

```

aacaaaaaac agtctgcgct gtcgactttc atcgagcgcc tctcttctgg tctgcgcatt 60
aacagcgcta aagatgacgc tgcggggccag gcgattgcta accggttccac ttctaacatc 120
aaaggtctga ctcaggccgc acgtaacgcc aacgacggta tctctctggc gcagaccact 180
gaaggcgcac tgtctgaaat caacaacaac ttgcagcgtg tctctgagct gaccgttcag 240
gccactacgc gtactaactc tgattctgac ctgtcttcaa tcaggacga aatcaaatcc 300
cgtctcgatg aaattgaccg cgtatccggt cagactcagt tcaacggcgt gaacgtactg 360
gcaaaagata acaccatgaa gattcaggtt ggtgcgaacg atggtcagac tatatccatc 420
gacctgcaaa aaatcgactc ttctactctt ggtttgaacg gtttctccgt ttctaaaaat 480
gctctcgaaa ctagcgaagc gatcactcag ttgccgaacg gtgcgaatgc accaatcgct 540
gtgaagatgg atgcgtctgt tctgaccgat cttaacatta ctgatgcttc cgctgtttcg 600
ctgcacaacg taactaaagg tgggtgcgca acgtctactt atgtgttcca gtatggcgat 660
aagagctatg cagcatctgt tgatgcggga ggtacagtaa aactgaataa agccgacgta 720
acataaacgc acgcagcaaa tgggtgttac aatgccaccg agattggtag tctggttcag 780
gttgggtgctg atgcaaacaa tgatgcagtt ggtttgttta ccgtgcaggg gaaaaactat 840
gttgctaatg actcattagt caatgctaat ggcgtctctg gcgctgcagc aactagagtt 900
acaattgatg ctgatggtag ccttggagct aaccaggcta aaattgaaat tagccaaaat 960
gggtgctactg ctgcaacatc agagttctgct ggtgcttcaa ccaacgatcc actgactctg 1020
ctggacaaag ctatgcgcat tgttgataaa ttccgttctt ctttgggggc ggtacagaac 1080
cgctgcagct ccgctgtaac caacctgaac aacaccata ccaacctgtc tgaagcgcgt 1140
tccgctattc aggacgcgca ctatgcgacc gaagtgtcca acatgtcgaa agccgagatc 1200
atccagcagg caggttaactc cgtgctgtcc aaa
1233

```

&lt;210&gt; 29

&lt;211&gt; 1713

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli



- 19 -

&lt;400&gt; 29

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60  
aaccagtcgt cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120  
gcgaaggatg acgcgcgagg tcaggcgatt gctaaccggt ttactttcaa cattaaggcg 180  
ctgactcagg ctgcacgttaa cgccaacgac ggtattttctg ttgcacagac cactgaaggc 240  
gcgctgtccg aaatacaaaa caacttacag cgtattccgtg aactgacggt tcaggcgacg 300  
accggaaacta actccacctc tgacctggac tccattcagg acgaaataca atcccgcttt 360  
gatgaaattg acccggtatc cggccaaacc cagtccaacg gcgtgaacgt actgtcaaaa 420  
gatggctcga tgaataatca ggtcggcgca aatgatgggt aaaccatcac gattgatctg 480  
aaaaagatcg actcttctac attgaagctg accagcttca atgttaacgg taaaggcgct 540  
gttgataatg ctaaagccac tgaagcagat ctgaccgctg cgggcttctc ccaaggtgca 600  
gtcgtcagtg gcaacagcac ctggactaaa tctactgtta ctacccttaa tgcagcaaca 660  
gctaccgacg tgctggcaag cgttagcggc ggcagcacta ttacgggtta taccggtaca 720  
aacaatggat taggcgtagc ggcttctact gcataacct acaacgcaac cagcaagtct 780  
tattcatttg acgcaaccgc acttaccat ggcgatggta ctggggccac cactaaagtt 840  
gctgatgtgc tgaagcccta tgcagcaaac ggtgataata cggctcagat ctccatcgcg 900  
ggaagcgctc aggcagttaa aattgccagc gatggcacc tgactgacgt caatgggtgat 960  
gctttatata ttggttctga cggcaacctg actaaaaacc aggcggcgcg tccagatgcg 1020  
gcaacgcttg acggtatttt caacggtgcg aatggtaatg cagcagttga tgcgaagatt 1080  
acattcggca gcggcatgac cgttgatttc acccaggcta gcaaaaaagt ggatattaag 1140  
ggcgcaacgg actccgcgga agatatggac actcgtctaa ctggggcagc ttataccgta 1200  
gctaacggcg cacagctctt tgacgttgcc gctggtgggg cagtaaacgc tactacaggt 1260  
ggcgctacgg taaatatttg tgctgatggg gaactgacga ctgcgaccaa caagatgtc 1320  
acagaactct atcagcaatt tgctaacggc aatattctgg atgatgcag cgcgctctg 1380  
tacaagcggg ctgacggttc tctgacctac gaagctactg gtaaatccga agtgaccacg 1440  
gatccgctga aagcgctgga cgtatctatc gcactcgtag acaaatccg ctccctccct 1500  
gggtcggtgc agaaccgtct ggattccgca gtcaccaacc tgaacaacac cactaccaac 1560  
ctgtctgaag cgcagtcctg cattcaggac gccgactatg cgaccgaagt gtccaatatg 1620  
tcgaaagcgc agatcatcca gcaggcgggt aactccgtgc tggcaaaagc caaccaggta 1680  
ccgcagcagg ttctgtctct gctgcagggt taa 1713

&lt;210&gt; 30

&lt;211&gt; 1668

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 30

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60  
aaccagtcgt cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120  
gctaaggatg acgcgcgagg tcaggcgatt gctaaccggt ttactttcaa cattaaggcg 180  
ctgactcagg ctgcacgttaa cgccaacgac ggtattttctg ttgcgacagc cactgaaggc 240  
gcgctgtccg aaatacaaaa caacttacag cgtattccgtg aactgacggt tcaggcttct 300  
accgggacta actccgattc ggatctggac tccattcagg acgaaataca atcccgctctg 360  
gacgaaattg acccggtatc tggccagacc cagtccaacg gcgtgaacgt actggcgaaa 420  
gacggttcaa tgaataatca ggttggtgag aatgagcgcc agactatcac tattgatctg 480  
aagaaaattg actcagatag cgtggggctg agtgggttta atgtgaatgg tggcggggct 540  
gttgctaata ctgcagcgac taaagatgat ttggtcgctg catcagtttc agctcgcgta 600

ggtaataaat acactgtctc tgcctggcctg tcgaaataca ctgctgctga tgttattgct 660  
 agtctcacag atggtgacgac agtaactgcy gctgggtgtaa gcaatgggtt tgcctgacgg 720  
 gcaactggag atgcttataa attcaatcaa gcaacaaca cttttactta caataccacc 780  
 tcaacacggc cagaactcca atcttaccctc acgcctaagg cggggggatc cgaacttttc 840  
 tccgttgaaa ttggtggcac caagcaggat gttgttctgg ctagtgtatg caaaactaca 900  
 gcaaaagacg ggtctaaact ttatattgac accacaggga atttaaccca aaacggttga 960  
 ggtactttag aagaagctac cctcaatggc ttacttttca accactctgg tccagccgct 1020  
 cgtgtacaat ctactattac tactcgggat ggaacttcaa tagttctatg aggtttctgg 1080  
 gacttttgaa caacaaaac tactggggct attaatgtca caggagcagt gatcagtct 1140  
 gatgcacttc tttcccgacg taaagcgact ggggttactt ctggcactta taccgtaggt 1200  
 acagatggag ttgttaaatc tgggtggcaat gacgtttata acaaagctga cgggacggga 1260  
 ttaactactg acaataccac aaaatattat ttacaagatg acgggtctgt aactaatggt 1320  
 tctggtaaag ctgtgtatgc tgatgcaaca ggaactata ctactgacgc tgaactataa 1380  
 gccgaacca ccgccgatcc cctgaaagct ctggacgaag cgtacgctc catcgacaaa 1440  
 tcccgcttct cctcctggtc ggtgcaaac cgtctggatt ccgcggtcac cactgtgacc 1500  
 aacaccacta ccaacctgtc cgaagcgcag tcccgattc aggaacgctga ctatgcccag 1560  
 gaagtgtcca acatgtcgaa agcgcagatc atccagcagg ccggttaactc cgtgctggca 1620  
 aaagctaacc aggtaccgca gcaggttctg tctctgctgc aggggttaa 1668

&lt;210&gt; 31

&lt;211&gt; 1713

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 31

atggcacaag tcattaatac caacagcctc tgcctgatca ctcaaaaata tatcaacaag 60  
 aaccagtcct cgtctgcgag ttctatcgag cgtctgtctt ctggcttctg tattaacagc 120  
 ctgacactcag ctgacgttaa ccgcaacgac ggtattttcc ttgcgcagac caccgaaggc 180  
 gcgctgtccg aaatcaacaa caacttacag cgtatccgtg aactgacggt tcaggccact 300  
 accggtaact actccgatc tgacctggac tccatccagg acgaaataca atctcgtctt 360  
 gatgaaattg acccgctatc tggctcagacc cagttcaatg cgtgtaattg ttgtccaaa 420  
 gacggttcaa tgaataattc ggtggcgcca aatgatgggt aaacatcac gattgacctg 480  
 aaaaaaatcg actctttctac actgaagctg accagcttca acgtcaacgg taaaggcgct 540  
 gttgataatg caaaaagccac tgaagcagat ctgacgcctg cgggcttctc ccaaagtcca 600  
 gttgtcagtg gcaatgacac ctggactaaa tctactgtta ctacctttaa tgcagcaaca 660  
 gctaccgatg tgcctggctg cgttagtgcc ggacgacta ttagcggtta tgcctggaca 720  
 aacaattgggt taggcgtagc ggcttctact gcataacct caacgcgaac cagcaagtct 780  
 tattcatttg acgcaaccgc acttactaat ggtgatggta ctgcggggctc aactaaagt 840  
 gctgatgttc tgaagacctc tgcagcaaac ggcgataaca cggctcagat ctccatcggt 900  
 ggtgacgctc aggaaggttaa aattgccagc gatgtgtacc tgacgggata taatggcgat 960  
 gctttatata ttggtctgta cggtaacctg acgaaaaacc aggcgcggcg cccagccgag 1020  
 gcaactgttg acggtatttt caacgggtgcg aatggctcatg atgcagttga tgcgaagatt 1080  
 accctcggca cggcgatgac cgttgacttc acccaggtta gcaacaatgt ggtatttaag 1140  
 ggcgcgacgg tatccgcgca agatatgaac actgcgttaa ccggtcagcg ttataccgta 1200  
 gctaacggcg cacagtctta tgcagttgcc gctgatgggt cagttaactc tactacaggt 1260  
 ggagcgacgg taatatattg tgcgtagggg gaactgacga ctgcggccaa caagactgtc 1320  
 acagaaactt atcacgaatt tgcatacggc aatatctggt atgatgacg cggcgctctg 1380

tataaagcgg ctgacggctc tctgaccact gaagctacag gtaaattcga agcgaccacg 1440  
 gatcccgctga aagcgctgga cgatgctatc gcattccgtag acaaattccg ttcttccctc 1500  
 ggtgcccgtgc agaaccgtct ggattccgca gtcaccaacc tgaacaacac cactaccaac 1560  
 ctgtccgaag cgcagctccg tattcaggac gcgactatg cgaccgaagt gtccaacatg 1620  
 tcgaagcgc agattattca gcaggcagggt aactccgtgc tggcaaaagc taaccaggta 1680  
 ccgcagcagg ttctgtctct gctgcagggt taa 1713

&lt;210&gt; 32

&lt;211&gt; 1188

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 32

aacaaaaacc agtctgcgct gtcgacttct atcgagcgcc tctcttctgg tctgcgcatt 60  
 aacagcgcta aagatgacgc tgcggggccag gcgattgcta accgcttca cttctaaccat 120  
 aaaggtctcga ctacaggccgc acgtaacgcc aacgacgcta tctctctggc gcagaccact 180  
 gaaggcgac tgcttgaat caacaacaac ttgcagcgtg tgcgtgagtt gactgttcag 240  
 gcgacgacgc ggactaactc tgattctgac ctgtcttcta ttcaggacga aatcaaatcc 300  
 cgtctgggat aaattgacgc tgtttccggt cagaccaggt tcaacggcgt gaactgtctg 360  
 gctaaaaacg gttctatggc gattcaggtt ggcgcgaatg atgggcgacac catcaacatc 420  
 gacctgcaga aaatcgactc ttctactctg ggctctggcg gcttctccgt atctaacaat 480  
 gcaactgaac tgagcgattc tatcactcag gttggtgcga gtggttcaat ggcagatgtg 540  
 aaactgagct ctgttgcttc ggctctgggt gttagcgcga gcaactctgac tctgcacaac 600  
 gtacagaccc cagctggcgc agcaacagct aactatgttg tctcttctgg ttctgcacaac 660  
 tactcagtat ctgttgaaga tagctccggt acagttaacg tgaacaccac tgatataggt 720  
 tataccgata ccgctaatgg cgttactacc ggttccatga ctggttaagta cgtttaaagt 780  
 ggagctgatg cattgggtgc tgcgttaggt tatgtcacgc tacagggaca aaacttcaaa 840  
 gctgatgctg gcgcgctggt taactccaag aatgctgctg gtatgcagaa gttacttctc 900  
 gcaattggcg atattgctaa taaagcgaat gctaaccatt acactggaa cttcttctgca 960  
 gatccactgg ctctgctgga caaagctatc gcattctgtg ataaattccg ttcttctcta 1020  
 gggggcggtc agaaccgtct gagctctgct gtaaccaacc tgaacaacac cactaccaac 1080  
 ctgtccgaag cgcagtcacc tattcaggac gccgactatg cgaccgaagt gtccaacatg 1140  
 tcgaaagcgc agatcatcca gcaggcgggt aactccgtgc tgtctaaa 1188

&lt;210&gt; 33

&lt;211&gt; 1638

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 33

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60  
 aaccagctct cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120  
 gcgaaggatg acgcccgcgg tcaggcgatt gctaaccggt ttactttctaa cattaaaggc 180  
 ctgactcagg ctgcacgtta cgccaatgac ggtattttctg ttgcacagac cactgaaggc 240  
 gcgctgtccg aaatcaacaa caacttacag cgtattctgt aactgacggt tcaggcttct 300  
 accgggacta actctgattc ggaactgtgac tccattcagg acgaaatcaa atcccgctc 360  
 gcgaaattg acccgctatc cggctcagacc cagttcaacg gcgtgaacgt acctggcaaa 420  
 gacgggttca tgaataatca ggttgggtgc aacgacggcc agactatcac tattgatctg 480

aagaaaattg actctgatac gctggggctg agtggggtta acgtaaatgg tagcgcagat 540  
 aaggcaagtg tcgcggcgac agctgacgga atgggttaaag accgatatat caaaggggtta 600  
 acttcacatg accggcagcac tgcataatac aaactacag caaatactgc agcaaaagga 660  
 tctgatatctc ttgcggcgct taagactggc gataaaatta ccgcaacagg tgcaaatagc 720  
 ctgtgtgata atgcgacatc gacaacttat acttataatg caaccagcaa taccttctcc 780  
 tatacggctg accgtgttaa ccaaacgaat gctgcagcaa atctcatacc tcgacgaggg 840  
 aaaacgacag ctgcatcagt tactattggt gggacagcac agaattgtaa tattgatgat 900  
 tcgggcaata ttaactcaag tgatggcgat caactttatc tggattcaac aggtaacctg 960  
 actaaaaaac aggcggcgaa cccgaaaaaa gcaacggttt ctgggcttct cggaaatacg 1020  
 gatgcgaaag gtactgtctg taaaacaacc atcaagacag agggctgggtg aacagttaca 1080  
 gctgaaggta atacaggtag tgtaaaaaat gaaggtgcta ctgtttcagc atctgcattt 1140  
 accggcattg catattccgc caacacccgt gggataactt atgctgttgc cgcaataaat 1200  
 actcaaaatg gtttccctgc gggggatgac ttaacccagg atgctcaaac tgtttcaacc 1260  
 tactactcgc aagccgatgg caccggtcac aatagcgcag gcaagaatac ctataaagac 1320  
 cgtgatgggtg tctacagcac agagaataaa acatcgaaag cgtccgatcc attggctgcg 1380  
 ctgacgacgc caatcagctc catcgacaaa ttccgttcat ccttgggtgc tatccagaac 1440  
 cgtctggatt ccgcggctac caacctgaac aacaccacta ccaacctgtc cgaagcgtag 1500  
 tcccgatttc aggacgcgca ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatc 1560  
 atccagcagg ccggttaact cgtgctggca aaagctaacc aggtaccgca gcaggttctg 1620  
 tctctgctgc agggctaa 1638

&lt;210&gt; 34

&lt;211&gt; 2145

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 34

aacaaatctc agtcttctct gagctccgcc attgaacgtc tctcttctgc cctgcgtatt 60  
 aacagtgcta aagatgacgc agcaggtcag gcgattgcta accgttttac agcaaatatt 120  
 aaaggtctga ctcaggcttc ccgtaacgcg aatgatggta tttctgttgc gcagaccact 180  
 gaaggtgcgc tgaatgaaat taacaacaac ctgcagcgtg tacgtgaact gactgttcag 240  
 gcaactaacg gtactaacct tgacagcgat ctttcttcta tccaggtgca aataactcaa 300  
 cgtctggaag aaattgaccg tgtatctgag caaactcagt ttaacggcgt gaagtcctt 360  
 gctgaaaaata atgaatgaa aattcagggt ggtgctaagt atgggtgaac catcactatc 420  
 aatctggcaa aaattgatgc gaaaactctc ggccctggagc gtttataat cgtatggcgcg 480  
 cagaaagcaa ctggcagtg cctgatttct aaatttaagc cgacaggtac tgataactat 540  
 gatgttggcg gtgatgcta tactgttaac gtatagatcg gagctgggta atgactccaa 600  
 ctattgata gtgtttttatg ttacagataa gcccgatgac ttgtctatgc agctccaccg 660  
 attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc agattcaggt 720  
 tatgcgcctc aattcgtcgc gtatatcgct tgcgtattac gtgcagcttt cccctcaggc 780  
 gggattcata cagcggccag ccataccgta tccatatcac cagctcaaa ggtgacagca 840  
 ggctcataag acgcccacgc gtgcgcatac tgcgttcacc gaatacgtgc gcaacaaccg 900  
 tcttccggag cctgtcatac ccgtaaaaa gccagcgtg gcgcgattta gccccgacat 960  
 agtcccactg ttgcctcatt tccgcgcaga cgatgacgtc actgcccgcc tgatgcgcgc 1020  
 aggttaccga ctccggcctg agttttttaa gtgacgtaaa atcggttga gcccacagcc 1080  
 cataatcgcg gcagttgcc ccgatccaac gccatctatg gccatatcaa tgattttctg 1140  
 gtgcgtaccg ggttgagaag ccgtgtgaagt gaactgcagt tgccatgttt tacggcagtg 1200  
 agagcagaga tagcgtgat gtcggcggtt gcttttgcgc ttaacgacca ccccgtagt 1260

agctgaacag gagggacagc t gatagaaac agaagccact ggagcaccctc aaaaacacca 1320  
 tcatacacta aatcagtaag ttggcagcat taccgaggag ctgttaaaga tactacaggg 1380  
 aatgatattt ttgttagtgc agcagatggg t cactgacaa ctaaatctga cacaacata 1440  
 gctgttacag ggattgatgc tacagcactc gcagcagcgg ctaagaataa agcacaagaat 1500  
 gataaattca cgtttaaattg agttgaattc acaacaacaa ctgcagcggg tggcaatggg 1560  
 aatgggtgat attctgcaga aattgatggg aagtcagtgta ctttactgtg gacagatgct 1620  
 gacaaaaaag cttctttgat tacagtgtag acagtttaca aaaaatagcgc tggcctttat 1680  
 acgacaacca aagttgataa caaggctgcc acactttccg atcttgatct caatgcagct 1740  
 aagaaaacag gaagcacgtt agttgttaac ggtgcaactt acgatgttaa tgcagatggg 1800  
 aaaacgataa cggagactgc ttcttgtaac aataaagta tgtatctgag caaatcagaa 1860  
 ggtggtgacc cgattctggt aaacgaagat gcagcaaaat cgttgcaate taccaccaac 1920  
 ccgctcgaaa ctatcgacaa agcatggctt aaagttgaca atctgcgttc tgacctcggt 1980  
 gcgatacaaa accgttttga ctctgctatc accaaccttg gcaacacgct aaacaacctg 2040  
 tcttctgccc gtacgcgat cgaagatgct gactacgcga ccgaagtgtc taacatgtct 2100  
 cgtgcgcaga tctgcacaa agcgggtacc tctgttctgg cgcag 2145

&lt;210&gt; 35

&lt;211&gt; 1587

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 35

aacaagaacc agtctgcgct gtgcagttct atcgagcgtc tgtctttctg cttgcgtatt 60  
 aacagcgcga aggatgacgc cgcaggtcag gcgattgcta accggtttac ttctaactatt 120  
 aaaggcctga ctcaggcgtc acgtaacgcg aacgacggta ttcttgttgc gcagaccacc 180  
 gaaggcgcgc tgtccgaat caacaacaac ttacagcgtg tgcgtgaact gaccgttcag 240  
 gcaaccaccc gtaccacact ccagctgtac ctggactcta tccaggacga aatataatcc 300  
 cgtctggacg aaattgaccg cgtatccggt cagaccagct tcaacggcgt gaacgtactg 360  
 gcaaaagacg gttccatgaa aattcaggtt ggcgcgaacg atggccagac catcacatc 420  
 gacctgaaga agattgactc ttctacgctg aaactgactg gttttaaact gaattggcaa 480  
 gcagcgggtg ataattgtaa agcgcaggat gcaaatctga ctaccgcggc ttttacacaa 540  
 ggcgttgttg attcaaatgg taatagtact tggactaaat caactacgac taatttcgat 600  
 gcggcaactg cagtaaacgt actagcagca gttaaagatg gcagcacaat caattacacc 660  
 ggtactggta atggttttag gattgctgca acaagtgett atacatatca cgtatgcaat 720  
 aaattctata cctttgatcc tacgggggct gcagtagctg gtgcccgcgc cagcctgcaa 780  
 ggtacttttg gtacagatag gaatactgca aaaatcacca tgcgtggttc tgctcaagaa 840  
 gtaaacatcg ctaaaagatg gaaaattact gatactgatg gtaaaagctt atatatcgat 900  
 tccactggta atttgactaa gaacggctct gatactttaa ctcaggcaac atggaatgat 960  
 gtccttactg gtgtaattc agttgatgat acaaggattg acttcgatag cggcatgtct 1020  
 gtcacccttg ataaagtga cagcactgta gatatactg gcgcacttat ttcagccgct 1080  
 gcaatgacta atgattgac aggtaaaggcc tataccgttag taaatgggtc agaattctac 1140  
 gctgtagcta ctaataacac agtaaaaacg actgctgatg ctaaaaatgt ttatgttgat 1200  
 gctagtggta aatatactac tgatgacaaa gccactgtta cagaaaacta tcatgaattt 1260  
 gcgaatggca atattctatg tgataaaggc gctgctgttt atgcggcggc ggatgttctc 1320  
 ctgactacag aaactacaag taaatcagaa gctacagcta acccgctggc cgtctctggac 1380  
 gacgcaatca gccagatcga caaatccgtt tcatccctgg gtgctatcca gaaccgtctg 1440  
 gattccgcag tcaccaacct gaaccaacac actaccaatc tgtctgaagc gcagtcocgt 1500  
 attcaggacg ccgactatgc gaccgaagtg tccaatatgt cgaaggcaga gatcatccag 1560

caggcaggca actccgtgct ggcaaaa

1587

&lt;210&gt; 36

&lt;211&gt; 1245

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 36

aacaaaaacc agtctgcgct gtcgaacttct atcgagcgcc tctcttctgg tctgcgcatt 60  
aacagcgcta aagatgacgc tgcggggccag gcgattgcta accgcttcac ttctaaccatc 120  
aaaggctctga ctacggccgc acgtaacgcc aacgacggta tctctctggc gcagaccact 180  
gaaggcgacac tgtctgaaat caacaacaac ttgcagcggt tctgtgaact gaccgttcagc 240  
gccactaccg gtactaactc tgattctgac ctgtcttcaa tccaggacga aatcaaatcc 300  
cgctctcgatg aaattgacgc cgtatccggt cagactcagt tcaacggcgt gaacgtactcg 360  
gcaaaagatg gctcgatgaa aattcagggt ggtgcaaatg atggtcagac aatcagcatt 420  
gatttgacga agattgattc ttctacttta gggttaaatg gttttctgtg ttccaaaaat 480  
gcagttactctg ttggtgatgc tattaactcaa ttgctggcg agacggcagc cgtgacacca 540  
gtaaccatca agtttgatga ttcagtaaaa actgatttaa aactgaccga tgcctcaggg 600  
ttaagtctgc ataactcaa agatgaaaat ggtaatttaa ctaaccagta tggttgacag 660  
aatggcgga aatcttaccg tgcacagtc gctgccaatg gtaatgttac gctgaaccaa 720  
gcaaatgtaa cctacagcga tgtcgcaaac ggtattgata ccgcaacgca gtcaggccag 780  
ttagttcagg ttggtgcaga ttctaccggt acgcaaaaag cattctgtgc tgtccaaggt 840  
aaaagctttg gcattgatga cgcgcctctg aagaataaca ctggtgatgc taccgtactc 900  
caaccgggaa catctgggac aacagttgtc gcagcgctcaa ttcatctgag tacggggcaa 960  
aactctgtag acgctgatgt aacggcttcc actgaattca cagggtgttc aaccaacgat 1020  
ccaactgact tgcctggaca agctatcgca tctgttgata aattccgttc ttctttgggg 1080  
gcggtacaga accgtctgag ctccgctgta accaacctga acaaacacac caccacactcg 1140  
tctgaagcgc agtcccgat tcaggacgac gactatcgca ccgaagtgtc caacatgtcg 1200  
aaagcgcaga ttatccagca ggcaggtaac tccgtgctgt ccaaa 1245

&lt;210&gt; 37

&lt;211&gt; 1185

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 37

aacaaaaacc agtctgcgct gtcgaacttct atcgagcgcc tctcttctgg tctgcgcatt 60  
aacagcgcta aagatgacgc tgcggggccag gcgattgcta accgcttcac ttctaaccatc 120  
aaaggctctga ctacggcgtc acgtaacgcc aatgacggta ttctcttagc acagacagcg 180  
gaaggcgacac tgtcagagat taacaacaac ttgcagcggt tgcgtgagtt gaccgttcagc 240  
gcaaccactg gtaccaactc tgattccgat ctctcttcta ttcaggatga aattaaatct 300  
cgctctggatg aaattgacgc cgtctctggt cagaccaggt tcaacggcgt gaacgtactcg 360  
gtcaaaaacg gttctatggc aattcagggt ggcgcgaacg atggccagac tatctctatc 420  
gacctgcaga aaatagactc ttctactctg ggtctgagcg gcttctctgt ttctcagaac 480  
tcctgaaac tgagcgattc tatcactacg atcggaata ctactgtgc actgaagaac 540  
gtggacctga gcgcagtagc aactaaactg ggcgtgaatg caagcaccct gagcctgcac 600  
gaagttcagg actctgctgt tgacgggtact ggtaccttcc ttgtttcttc tggcagcgac 660  
aactatgtct gttctgtaga cgcgcctctc ggtgcagtta actgaacac cactgacgtc 720

acctatgatg acgctactaa tgggtgttact ggcgcgactc agaacgggtca gctgatcaaa 780  
 gtaactttctg acgccaacgg tgcagctgtt ggttacgttaa ccattcaggg taaaaactat 840  
 cagggtcgtg cgaccgggtg tgacgtttctg gcgaacagcg gtgttgacgc tccaactaca 900  
 ctgcttgata ccggtactct gcaactgagc ggtagctggg caactactga gctgaaaggt 960  
 actgcaactc agaaccacct ggcaactatt gacaaagcta tgccttctgt tgataaattc 1020  
 cgtttctctc tgggtgcggg acagaatcgt ctgagctctg ctgtaaccaa cctgataaac 1080  
 accaccacta acctgtctga agcgcagctc cgtattcagg atgccgacta tgcgaccgaa 1140  
 gtgtcaataa tgtctaaagc gcagatcgtt cagcaggccg gtaac 1185

&lt;210&gt; 38

&lt;211&gt; 1383

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 38

aacaaatctc agtctctctc tagctctgct attgagcgct tgtctctctg tctgcgtatt 60  
 aacagcgcaa aagacgatgc agcagggtcag gcgattgcta accgttttac ggcaaatatt 120  
 aaaggtctga ccagggtctc ccgtaacgca aatgatggta ttctgtgtgc gcgaccact 180  
 gaaggtgcgc tgaatgaaat taacaacaac ctgcagcgta tctctgaact ttctgttcag 240  
 gcaactaacg gtactaactc tgacagcgat cttctctcta tccaggctga aattactcaa 300  
 cgctcggaaag aaattgaccg tgtatctgag caaactcagt ttaacggcgt gaaagtctct 360  
 gctgaaaata atgaaatgaa aattcagggtt ggtgctaata atgggtgaaac catcactatc 420  
 aatctggcaa aaattgatgc gaaaactctc ggccctggag gttttaatat cgatggcgcg 480  
 cagaaagcaa caggcagtgaa cctgatttct aaattttaaag cgacaggtaac tgataattat 540  
 gatgttgccg gtaaaactta taccgtgaat gtggagagcg gcgcgggttaa gaatgatgct 600  
 aataaagatg tttttgtaag cgcagctgat ggatcgctga cgaccagtag tgatactaaa 660  
 gtatccggtg aaagtattga tgcaacagaa ctagcgaaac ttgcaataaa attagctgac 720  
 aaaggctcca ttgaatacaa gggcattaca tttactaaca acactggcgc agagcttgat 780  
 gctaattgta aaggtgtttt gaccgcaaat attgatggtc aagatgttca attactatt 840  
 gacagtaatg caccacggg tgccggcgca acaataacta cagacacagc tgtttacaaa 900  
 aacagtgcgg gccagttcac cactacaaaa gtggaaaaa aagccgcaac actctctgat 960  
 ctggatctta atgcagccaa gaaaacaggt agcactttag ttgtaaatgg cgccacctac 1020  
 aatgtcacgc cagatggtaa aacggttaact gatactactc ctggtgcccc taaagtgatg 1080  
 tatctgagca aatcagaagg tggtagcccg attctggtaa acgaagatgc agcaaatgc 1140  
 ttgcaatcta ccaccaaccc gctcgaaact atcgacaagg cattggctaa agttgacaa 1200  
 ctgcgttctg acctcgtgtg agtacaaaac cgtttcgact ctgccatcac caaccttgcc 1260  
 aacacggtaa acaacctgtc tctgcgccgt agccgtatgc aagatgctga ctacgcgacc 1320  
 gaagtgtcta acatgtctcg tgcgcagatc ctgcaacaag cgggtacctc tgttctggcg 1380  
 cag 1383

&lt;210&gt; 39

&lt;211&gt; 1680

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 39

atggcacaag tcaataatac caacagcctc tcgctgatca ctcaaaaataa tatcaacaag 60  
 aaccagtcgt cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120

gcgaaggatg acgccgcagg tcaggcgatt gctaaccgtt tcacctctaa cattaaaggc 180  
 ctgactcagg ctgcacgtaa gcgcaacgac ggtatttctg ttgcacagac caccgaaggc 240  
 gcgctgtccg aaatcaacaa caacttacag cgtatccgtg aactgacggt tcaggcttct 300  
 accgggacta actctgattc ggatctggac tccattcagg acgaaatcaa atccgctctg 360  
 gacgaaattg accgcgtatc cgccagacc cagttcaacg gcgtgaacgt gctggcgaaa 420  
 gacggttcaa tgaaaaattca ggttggtgcg aatgacggcc agactatcac tattgatctg 480  
 aagaaaaatt actctgatac tctgggtttg agtggattta atgtgaatgg caaaggggct 540  
 gtggctaacy caaaagcgac cgaagcagat ttaacggggg ctggtttctc tcaaggagcg 600  
 gtggatacaa acggaaatag tacttggaca aaatcaacca ccaccaatta ctgagctgca 660  
 acaactgctg acttggtatc gaccattaag gatggctcta ctgttataca tgcagggaca 720  
 gacaccggat taggggtcgc agcagcagga aattatactt atgatgcgaa cagtaaatct 780  
 tattccttca atgccaattg tctgacgggc gcaaataccg caactgcact caaaggttac 840  
 ttggggacag gtgctaacac cgctaaaatt tctatcggtg gtacagagca ggaagtgaat 900  
 attgccaag atggcactat tacagatacg aatgggtgat cgctctatct ggaattacc 960  
 ggcaacctga ctaagaacta tgcgggttca ccactgcag caacgctgga taacgtatta 1020  
 gcttcgcaaa ctgtaaatgc cactatcaag tttgatagcg gtatgacggt tgattacact 1080  
 gcaggctact gcgcgaatat tacagggtga tccatttctg cagatgacat ggcgcgaaaa 1140  
 ctgagcggaa aggcgtacac tgttgccaat ggtgctgagt cttatgacgt tgctgcagtt 1200  
 acggggcgctg taacaactac agcaggttaat tcacctgtgt atgccgatgc agacggtaaa 1260  
 ttaacgacga tgcgcagtaa tacgggtact cagacttacc acgagtttgc taatggtaac 1320  
 atttatgatg acaaggctc gtcactgtat aaagctgcag atggctctct gacttctgaa 1380  
 gctaaaggga aatctgaagc aaccggcgat cccctgaaag ctctggagca agccatcagc 1440  
 tccatcgaca aattccgctc ctccctcggt gccgttcaaa accgcttgga tctcgcggtg 1500  
 accaactgga acaacaccac taccaacctg tctgaagcgc agtcccgtat tcaggagccc 1560  
 gactatgcga ccgaagtgtc caatatgtcg aaagcgacga tcatccagca ggcgggtaac 1620  
 tccgtgttgg caaaagctaa ccaggatacc cagcaggttc tgtctctgct gcagggttaa 1680

&lt;210&gt; 40

&lt;211&gt; 1146

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 40

gcgctgtcga ettcattcga gcgctctctc tctggtttgc gattaacag cgctaaagat 60  
 gacgctgcgg gccaggcgat tgcataaccg ttacttcta acatcaagg tctgactcag 120  
 gcgcacgta acgccaacga cggatctctc ctggcgacga ccaactgaag cgaactgtct 180  
 gaaatcaaca acaacttgca gcgtgttctg gaactgaccg ttcaggccac taccggtact 240  
 aactctgatt ctgacctgtc ttaactccag gacgaaatca aatcccgtct ggctgaaatc 300  
 gatcggtctc ctggctacac ccagttcaac ggctgaaacg tgctggctaa aaacggttct 360  
 ctgaattatc aggttggcgc gaatgatggg cagaccatct ctatcgattt gcagaaaata 420  
 gactctctcg cctctggttt aagtgggttt agtgttgccg gtggggcgct aaaattaaagc 480  
 gatacagtag ccgaggtcgg cgatggttca gcgcgcacg ttaaagtga ctctggtatga 540  
 gcagcaacag atattggtac tgctttgggg caaaaggtta atgcaagttc ttaacggttg 600  
 cacaatatct tagacaaaga tggtgccgca actgagaact atgtgtgttag ctatggtagt 660  
 gataattacg ctgcaactgt tgcatgatgc gggactgtaa ctcttaataa acgggatact 720  
 acttattcag gcggtgatat taccggcgct accaaagatg atacgttgat taaagttgct 780  
 gctaattctg acggagaggc cgttggttct gctaccgttc agggtaagaa ttatgaaatt 840  
 acagatgggtg taaaaaacca gtccactgct gcaccaaccg atattgctca gaccattgat 900



ctggatcagg ctgatgaatt tactggggct tccactgctg atccactggc accttttagac 960  
 aaagctattg cacaggttga tactttccgc tctccctcg gtgcggttca aaacgctcg 1020  
 gattccgcag tcaaccaact gaacaacact actaccaacc tgtctgaagc gcagtcgctg 1080  
 attcaggacg cgcactatgc gaccgaagtg tccaatatgt cgaagcgca gatcatccag 1140  
 caggcc 1146

&lt;210&gt; 41

&lt;211&gt; 1506

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 41

atggcacaa gtcattaatc caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60  
 aaccagtcgt cgctgtcgag ttctatcgag cgtctgtctt ctggcttggc tattaacagc 120  
 gcgaaggatg acgacggggg tcaggcgatt tctaacgctg ttactcttaa tattaaaggc 180  
 ctgactcagg ctgcacgtaa cgccaatgac ggtatttctc tggcgagac cactgaaggc 240  
 gcaactgtct aaatcaacaa caacttgcag cgtgtgcgtg aactgacgct acaggcgaca 300  
 accggaacga actccgaatc tgacctgtcc tctatccagg acgaaatcaa atccgctctg 360  
 gaagagattg accgcgtatc cggccagact cagttcaacg gcgtgaatgt gctggcaaaa 420  
 gacggcacca tgaanaattca ggtaggcgcg aacgatggtc agactatctc tatcgatctg 480  
 aaaaaaatcg actcttcaac cctgggcctg accggttttg atgtttcgac gaaagcgaat 540  
 atttctacga cagcagtaac gggggcgcca acgaccactt atgctgatag cgccgttgca 600  
 attgatattc gaacggatat tagcgggtatt gctgctgatg ctgcggttag aacgatcaat 660  
 ttcgataata caacaggcaa gttactacga cagattacca gtgcggccaa tccgggcctt 720  
 gatggtgctt atgaatcca tgtaaatgac gcggatggtt ccttcaactg agcagcgagt 780  
 gataaacaag cgggtgtctg tccgggtact gctctgacaa gcggtaaaat tcagactgca 840  
 accaccacgc caggtacggc tgttgatgct actgcggcta aaactgctct ggtcgagca 900  
 ggtgctgaca cgagtggcct gaaactggtt caactgtcca acacggattc cgcaggtaaa 960  
 gtgaccaacg tgggttacgg cctgcagaat gacagcgcca ctatctttgc aaccgactac 1020  
 gatggcacca ctgtgaccac gccggcgcca gagactgtga cttacaaga tgcttcgggt 1080  
 aacagcacca ctgcggctgt cacactgggt ggctctgatg gcaaaaccaa tctggttacc 1140  
 gcgctgacg gcaaaacgta cgggtgcgact gcactgaatg gtgctgatct gtccgatctc 1200  
 aataaacacg ttaaatctgt tgcagacaac gctaaacggt tggtgcctct ggatgatgca 1260  
 attgcgatg tgcacaatt ccgctctctc ctccggtgcg tgcaaaacgg tctggattcc 1320  
 gcagtcacca acctgaacaa caccactacc aacctgtctg aagcgagtc cgttattcag 1380  
 gacggcgact atgcgacgca agtgtccaac atgtcgaaag cgcagattat cagcagcgga 1440  
 ggtaactccg tgctgtccaa agctaaccag gttccgcagc aggttctgtc tctgctgcag 1500  
 gggttaa 1506

&lt;210&gt; 42

&lt;211&gt; 950

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 42

aacaaaaacc agtctgcgct gtcgacttct atcgagcgcc tctctcttgg tctgcgtatt 60  
 aacagcgcta aagatgacgc cgcgggccag gcgattgcta accgctttac ttctaacatc 120  
 aaaggtctga ctacggccgc acgtaaccgc aacgacggta tttctcttgc gcagacggct 180

gaaggcgcgc tgtcagagat taacaacaac ttgcagcgta ttctgtaact gaccgttcag 240  
 gctcttaccg gcacgaacte tgattccgac ctgtcttcta ttccaggaga aatcaaatcc 300  
 cgtcttgatg aaattgaccg tgtatctggt cagaccaggt tcaacggtgt gaacgtgctg 360  
 tcgaaaaacg attcgatgaa gattcagatt ggtgccaatg ataaccagac gatcagcatt 420  
 ggcttgcaac aatcgacag taccactttg aatctgaaag gattttaccgt tccggcatg 480  
 gcggaattca gcgcggcgaa actgacggct gctgatggta cagcaattgc tgctgcggat 540  
 tccaaggatg ctgggggttaa acaagtcatt ttactgtctt acactgacac cgcgtctaac 600  
 agtataaat atgcggtcgt tgattctgca accggttaaat acatgggaagc cactgtagcc 660  
 attaccggta cggcggcgcc ggtaactgtt ggtgcagcgg aagtggcggg agccgctaca 720  
 gcgcatccgt taaaagcact ggatgcgcga atcgctaagc tcgacaattt ccgctctccc 780  
 ctccggtgcc ttcaaaaccg tctggattct cgcggtacca acctgaacaa caccaccacc 840  
 aacctgtctg aagcgcagtc ccgtattcag gacgccgact atgcgaccga agtgtccaac 900  
 atgtcgaaag cgcagattat ccagcaggcc ggtaactccg tgctggcaaa 950

&lt;210&gt; 43

&lt;211&gt; 1707

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 43

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60  
 aaccagctct cgtctgcgag ttctatcgag cgtctgtctt ctggcttgccg tattaacagc 120  
 gcgaaggatg acgcagcggg tcaggcgatt gctaaccggt ttacctctaa cattaaaggt 180  
 ctgactcagg ctgcacgtta ccgcaacgac gctatttctg ttgcacagac gactgaaggc 240  
 gcgctgtccg aaatcaacaa caacttacag cgtatccgtg aactgacggt tcaggcttct 300  
 accgggacta actccgattc ggtctgtgac tccattcagg acgaaatcaa atcccgtctc 360  
 gacgaaattg accgcgtatc cgttcaaac ccagttcaacg gtgtgaacgt actggcgaaa 420  
 gacggttcga tgaanaattca ggttggtgag aatgacggcc agactatcac gattgatctg 480  
 aagaaaattg actcagatag cgtggggctg aatggtttca acgttaattg caaaggcact 540  
 attgcgaaca aagctgtctac agtcagcgat ctgacgcgtg ctgggtgcaac gggaacaggt 600  
 ccttatgctg tgaccacaaa caatacagca ctacgcgcta gcgatgcact gtctgcgctg 660  
 aaaacgggag atacagttac tactactggc tcgagtgctg cgactctata ttatgatgag 720  
 gctaaaggga acttcaaccac tcaagcaaca gttgcagatg gcgatgttgt taactttgcg 780  
 aatactctga aaccagcggc tggcactact gcacagagtg tttatactcg tagtactggt 840  
 gatgtgaagt ttgatgtaga tgctaattgc gatgtgacca tcggtggtaa agccgctgac 900  
 ctggaagcca ctggtaacct atctacaac aaccgcgca ttgcattctc actggaattg 960  
 tcogactctg ttgctagcgg tagtacetta gcgacaactg gttctatcca gctgtctggc 1020  
 acaactata actttggtgc agcggcaact tctggcgtaa cctacaccaa aactgtaagc 1080  
 ctgtgatact tactgagcac agtcgagagt gctgcgaact caacacagc agtactggtt 1140  
 gcgacaatta agtataatac aggtattcag tctgcaacgg cgtccttcgg tgggtgtaat 1200  
 actaatggtg ctggtgaattc gaatgacacc tatactgatg cagacaagaa gctcaccaca 1260  
 accgactctt acactatcaa ctacaacgct gataaggata ccggtacagt aacttagtgt 1320  
 tcaaatggcg caggtgcaac tggtaaaatt gcagctactg ttggggcaca ggccttatgt 1380  
 aactctacag gcaaacctgac cactgaaacc accagtcgag gcaactgcaac caaagatcct 1440  
 ctgggtgccc tggatgaagc tatcagctcc atcgacaaa tccggttcac cctgggtggt 1500  
 atccagaacc gtctggattc cgcggttaac aactgcaaca acaccactac caactgtctc 1560  
 gaagcgcagt ccgctattca ggacgccgac tatgcgacgg aagtgtccaa catgtcgaaa 1620  
 gcgcagatta tccagcaggc cggtaactcc gtgctggcaa aagccaacca ggtaccgcag 1680

caggttctgt ctctgctgca ggggttaa

1707

&lt;210&gt; 44

&lt;211&gt; 1720

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 44

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaata tatcaacaag 60  
 aaccagctctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120  
 gcgaaggatg acgcccaggg tcaggcgatt gctaaccggt ttacttctaa tattaaggcg 180  
 ctgactcagg ctgcacgtaa cgccaatgac ggtatttctg ttgcacagac cactgaaggc 240  
 gcgctgtcgc aaatcaacaa caacttacag cgtgtgcgtg aactgaccgt tcaggcgacc 300  
 accggtacca actccagtc tgatctggac ttatccagg acgaaatcaa atccgctctg 360  
 gacgaaattg accgcgtatc cggtcagact cagtccaacg gcgtgaacgt actggcaaaa 420  
 gacggttcca tgaaattcca ggttggcgcg aatgatggcc agaccatcac tatcgacctg 480  
 aagaagattg actcttctac gttgaaactg actggtttta acgtgaatgt tctgggttct 540  
 ttggcgaaata ctgcggcgac taaagacgaa ctggctgctg ctgctgcggc ggcgggtaca 600  
 actctctgtg tcggactga cgcgctgacc aaataaccg tagacgcagg gcttaacaaa 660  
 gccacagcag caaacgtgtt tgcaaacctt gcagatgggt ctggtgttga tgctagcatt 720  
 tccaacggtt ttggtgcagc agcagccaca gactacacct acaataagc tacaatgat 780  
 ttacttttca atgccagcat tgctgctggt gctgcggcgc gtgatagtaa cagcgagct 840  
 ctgcaatcct tctgtactcc aaaagcaggt gatacagcta acctgagcgt caaaatcggt 900  
 acgacatctg ttaattgtgt tctggcgagc gatggcaaaa ttacagcgaa agatggctca 960  
 gctctgtata tcgactcaac gggtaacctg actcagaaca gcgcaggcac tgtaacgaca 1020  
 gcaaccctgg atggactgac caaaaaccat gatgcgacag gagctgttgg tgttgatac 1080  
 acgaccgcag atggcgcaac tatctctctg gcaggctctg ctaacgcggc aacaggtact 1140  
 caatcagggt caattacact gaaaatgtt cgtatcagtg ctgatgctct gcagctctgt 1200  
 gcgaaggta ctggtatcaa tgttgataat ggtgctgatg atatttctgt tagtaaaacc 1260  
 ggtgtgtcgt actaccggag gtgcgcctac ttatactgat gctgatgta aattaacgac 1320  
 aaccaacacc gttgattatt tctgcgaac tgatggcagc gtaaccaatg gttctggtaa 1380  
 aggggtttac accgatgcag ctggttaaatt cactaccgac gctgcaacca aagccgcaac 1440  
 caccaccgat ccgctgaaag cccttgatga cgcatacgc cagatcgata agtccgttc 1500  
 atccctgggt gctatccaga accgtctgga ttccgcgggt accaacctga acaaccacc 1560  
 taccacacct tcggaagcgc agtcccgtat tcaggacgac gactatgcga cgaagtgtc 1620  
 caatagtctg aaagcgcaga tcattccaga ggcgcgtaac tccgtgttgg caaaagctaa 1680  
 ccaggtaccg cagcaggctc tgtctctgct gcagggttaa 1720

&lt;210&gt; 45

&lt;211&gt; 14516

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 45

gatctgatgg ccgtaggcgc ctacgtgctt tctgctgata tctgggctga gttggaaaaa 60  
 actgctccag gtgcctgggg acgtattcaa ctgactgatg ctattgcaga gttggctaaa 120  
 aaacagctctg ttgatgccat gctgatgacc ggcgacagct acgactgcgg taagaagatg 180  
 ggctatatgc aggcattcgt taagtatggg ctgcgcaacc ttaagaagg ggcaagttc 240

cgtaagagca tcaagaagct actgagttag tagagattta cactgttttg tgacgataag 300  
 ccagaaaaaa tagcggcagt taacatccag gcttctatgc ttttaagcaat ggaatgttac 360  
 tgccggtttt tatgaaaaat gaccaataat aacaaggttaa cctaccaagt ttaatctcgt 420  
 ttttgttgga ttttttcttg tttctggtcg catttggtaa gacaattagc gtgagtttta 480  
 gagagttttg cgggactctcg cggaactgct cacactcttg gcattttagt agtgcactggt 540  
 tagctgttaa gccagggggcg gttagcttgcc taattaattt ttaactgtata catttattct 600  
 tggcgcttat agcaataaaa gtcaatcgga ttaactttct tttccattag gtaaaagagt 660  
 gtcttagtcg gctcaggga atgtgttttg gttagtagtac ttttcaaat atccattttc 720  
 cgatttagat gccagttgat gttactatgc tgcatacata tcaatgtata ttatttactt 780  
 ttagaattgt atatgaaaa aatagtgtac ataggcaatg tagcgtcaat gatgttaagg 840  
 ttcaggaaag aattaatcat gaatttagtg aggcaaggtg ataattgtata ttgtctagca 900  
 aatgattttt ccaactgaaga tcttaaaagta cttctgctat ggggcgttaa gggggttaa 960  
 ttctctctta actcaaaagg tattaatctt ttttaaggata taattgtctg ttatgaacta 1020  
 aaaaaaatc ttaaggatat ttcccagat atgtattttt catattttgt aaagccagta 1080  
 atatttgga cttattgttc aaagtgtgca aaagtgcaca ggattgttgg aatgattgaa 1140  
 ggtctaggta atgccttcac tattataaag ggaagcaga ccacaaaaac taaaatgata 1200  
 aagtggtata aaattctttt atataagta gcattaccga tgcttgatga ttgtattcta 1260  
 ttaaatcatg atgataaaaa agatttaate gatcagtata atattaaagc taaggttaaca 1320  
 gtgttaggtg ggtattgatt ggtacttaat gagttttcat ataaagagcc accgaaagag 1380  
 aaaaattact ttatttttat agcaaggtta ttaagagaga aagggtattt tgagtttatt 1440  
 gaagccgcaa agttcgttaa gacaacttat ccaagttctg aatttgtaat tttagggagt 1500  
 tttagagata ataatctctt ctcatcaca aaaaatgaaa ttgaatcgct aagaaagaa 1560  
 catgattcta ttatctctgg tcatgtggaa aatgttcaag attgggttag aaaaagtctt 1620  
 gtttttgttt taactacatc atatcgagaa ggcgtaccac aggtgtacca agaagctatg 1680  
 gctatttgta gaectgtaat aacaactaat gtacctgggt gttaggataa aataatgat 1740  
 ggggtcaatg gcttttgat acctccattt gaaattaatt tactggcaga aaaaatgaaa 1800  
 tattttattg agaataaaga taaagttact gaaatggggc ttgctggaag gaagtttgca 1860  
 gaaaaaaact tttagtcttt tgaaaaaaat aatagactag catcaataat aaaaatcaat 1920  
 aatgattttt gacttgagca gaaattattt atatttcaat ctgaaaaata aaggctgtta 1980  
 ttatgaataa agtgcattta attactggtta tcaactggga agatggctcc tatttggcag 2040  
 aattatttgt agaaaaaggt tatgaagttc atggtattaa acgccgtgca tcttcattta 2100  
 atactgagcg agtggatcac atctatcagg attcacattt agctaatactt aaactttttc 2160  
 tacactatgg cgatttgaca gatacttoca atctgaccg tattttaaaa gaagttcaac 2220  
 cagatgaagt ttacaatttg gggcgatga gccatgtagc ggtatcattt gagtaccag 2280  
 aatacactgc tgatgttgat gcgataggaa cattgcgtct tcttggaagt atcagatat 2340  
 ttgggctgga aaaaaagaca aaattttatc aggcctcaac ttcagagctt tatggttttg 2400  
 ttcaagaaat tccacaaaaa gagactacgc cattttatcc acgttgcctt tatgctgttg 2460  
 caaaattata tgctatttg atcactgtta attactgtga gtcttattgt atgttgcctt 2520  
 gcaatggtag tctctttaac cagcaatcac ctgcgcgttg cgagaccttt gttactcgta 2580  
 aaataacacg cgggatagca aatattgtct aaggctctga taaatgttta tacttgggaa 2640  
 atatggattc tctgcgtgat tggggacatg ctgaaggata tgtcaaaagt caatggatga 2700  
 tgcgcagca agaaactcca gaagattttg taattgtctac aggaattcaa tattctgtcc 2760  
 gtgagtttgt cacaatggcg gcagagcaag taggcataga gttagcattt gaaggtgagg 2820  
 gagttaatat aaaaaggtgt gttgtttcgg tcaatggcac tgatgctaaa gctgtaaac 2880  
 cggcgatgt aatttatctt gtatagccaa ggtattttag gctgcagaa ttgaaacct 2940  
 tgcttggcga tctactaat gcgcataaaa aattaggatg gagccctgaa attacattgt 3000  
 gtgaattgtt aaaaagaatg gtttccagcg atttagcaat agcgaaaaag aacgtcttgc 3060  
 tgaaagctaa taacattgac actaatatc cgcaagaata aaaaagataa tacattaaat 3120

aattaaaaat ggtgctagat ttattagtagc cattattttt ttttgggtga ctaattgttta 3180  
ttacatccaga taaatttaga gaaattatca agtttagttcc attagtagtca attgatctgc 3240  
taattgaaaa cgagaaatggt gaattatttat ttggtctctag gaataatcga ccggccaaaa 3300  
attatttttt ttctccaggt ggtaggattc gcaaaaaatga atctattaaa aatgctttta 3360  
aaagaatatc atctatggaa ttaggttaag agtatggtat ttcaggaaagt gtttttaagt 3420  
gtgtatggga acattttctat gatgatgggt tttttcttga aggcgaggca acacattata 3480  
tagtgctttg ttacacactg aaagtcttta aaagtgaatt gaattctccca gatgatcaac 3540  
atcgtgaata cctttggcta actaaacacc aaataaatgc taacaagaat gttcataact 3600  
attcaaaaaa ttattttttg taatttttat taaaaattaa tatgcgagag aattgtatgt 3660  
ctcaatgtct ttaccctgtg attattgcgg gaggaacccg aagccgtctc tggccgttgt 3720  
ctcgtgatt ataccctaaa caatttttaa atttagttgg ggatcttaca atgttgcaaa 3780  
caacaattac gcgtttggat ggcacatgaat gcgaaaatcc aattgttatc tgcattgaag 3840  
atcccgatt tattgtagca gagcaattac gacagattgg taagctaacc aagaatatta 3900  
tacttgagcc gaaggccgt aactactgac ctgccatagc tttagctgct tttatcgctc 3960  
agaagaataa tctaatgac gaccctttat ttagtagtact tgcggcagac cactctataa 4020  
ataatgaaaa agcatttcga gagtcaataa taaaagctat ccgctatgca acttctggga 4080  
agttagtaac atttggaaatt attccggaca cggcaaatc ttggttatgga tatattaaga 4140  
gaagtcttct agctgatcct aataaagaat tccagcata taatgttgcg gagtttgttag 4200  
aaaaaccaga tgttaaaaaa gccacggaaat atatttcgag tgggaattat tactggaata 4260  
cggaatggt ttattttcgc gccagtaaat atcttgatga actacggaaa tttagaccag 4320  
atatttatca tagctgtgaa tgtgcaaccg ctacagcaaa tatagatatg gactttgttc 4380  
gaattaacga ggcgtgagtt attaatgttc ctgaagagtc tatcgattat gctgtgatgg 4440  
aaaaaacaaa agacgctgta gttcttccga tagatatggg ctggaatgag ctgggttctt 4500  
ggctcatcact ttgggatata agccaaaagg attgccatgg taatgtgtgc catggggatg 4560  
tgctcaatca tgatggagaa aatagtttta ttactctga gtcaagtctg gttgcgacg 4620  
tcgggtagaa taatttagta attgtccaaa ccaaggatgc tgtactggtt cgggaccgtg 4680  
ataaagtcga aatgtttaa aacatagtgg acgatctaaa aaagagaaaa cgtgctgaat 4740  
actacatgca tctgtgagtt ttctgccctt ggggtaaat cgtatgaata gaccaaggcg 4800  
atagatatag agtaaaaaaa ataagtatta aaccaggaga agggtttagat ttaaggatgc 4860  
atcatcatag ggcagagcat tggattgttg tatccgtac tgctaagatt tcaataggta 4920  
gtgaagttaa actattagtt tctaagtgt ctatatatc cctcaggga gcaaaaata 4980  
gtcttgagaa tccaggcgta ataccttgc atctaattga agtaagtctt ggtgattacc 5040  
ttgaatcaga tgatatagtg cgttttactg acagataata cagtaaacaa ttctcaaaagc 5100  
gagattgata aatatgaata aaataacttg cttcaaaagca tatgatatac tggggcgtct 5160  
tgggtgctgaa ttgaatgatg aaatagcata tagaattggt cgcgcttatg gtgagttttt 5220  
taaacctcaa actgtagtgg tgggaggaga tgctcgctta acaagtgaga gtttaagaaa 5280  
atcactctca aatgggctat gtgatgcagg cgtaaatgtc tttagatctt gaatgtgtgg 5340  
tactgaagag atatttttt ccacttggtta tttaggaatt gatggtggaa tcgaggtaac 5400  
tgcaagccat aatccaattg attataatgg aatgaaatta gtaaccaaag gtgctcgacc 5460  
aatcagcagt gacacaggtc tcaaaagat atcaacaatta gttagatata ataattttga 5520  
agagctcaac ctagaaaaaa aagggaatat taccaaatat tccaccgcag atgcctacat 5580  
aaatcatttg atgggctatg ctaactctgca aaaaaataaa aaatcaaaa tagtttgtaa 5640  
ttctgggaat ggtgcagctg gtctcgttat tgatgctatt gaggaatgct ttctacggaa 5700  
caatattccg attcagttg taaaaataaa taatacaccg gatggtaatt ttccacatgg 5760  
tatccctaat ccattactac ctgagtgcaag agaagatacc agcagtgccg ttataagaca 5820  
tagtgtgat ttgtgtattg catttgatgg tgatttttgg aggtgttttt tctttgatga 5880  
aaatggacaa ttattgaag gatactacat tgttggttta ttacggaaag tttttttagg 5940  
gaaatatcca aacgcaaaaa tcatcatga tcttcgctct atatggaata ctattgatat 6000

cgtagaaagt catggttgga tacctataat gactaaaacc ggctatgctt acattaaagca 6060  
 aagaatgcgt gaagaggatg ccgtatatgg cggcgaatg agtgcgcctc attattttaa 6120  
 agattttgca tactgcgata gtggaatgat tcttggatt ttaattttgt aactttttgag 6180  
 ctctgacaaat aaaaaattag gtgaactggg ttgtggtgt ttaaacgact ggcgcggcga 6240  
 tggagaaata aactgtacac tagacaatcc gcaaaatgaa atagataaat tatttaactg 6300  
 ttacaagat agtgccttag ctgttgatta cactgatgga ttaactatgg agttctctga 6360  
 ttggcgcttt aatggttagat gctcaaatat agaacctgtg tacgatgta atgtagaatc 6420  
 taggaataat gctattctta tgcaggaaaa aacagaaaga attctgaatt ttatatcaa 6480  
 ataaatttgc acctgagttc ataattggga caagaaatat atgaaagtac tcttgactgg 6540  
 ctcaactggc atggttggtg agaataat attagcatgat agtgcaagta aatataaat 6600  
 acttactcca accagctctg atttgaattt attagataaa aatgaaatag aaaaattcat 6660  
 gcttatcaac atgcagact gtattatata tgcagcggga ttagtgggag gcaattcatc 6720  
 aataataagc aggccgtttg attttctgga aaaaaattg cagatgggtt taaatttgt 6780  
 ttccgtcgca aaaaaactag gtatcaagaa agtgcctaac ttgggtagtt catgcatga 6840  
 ccccaaaac tttgaaggag ctattctctg gaaagctctg ttaactgggt agctagaaga 6900  
 aactaatgag ggaattgcta ttgcgaaat tgctgtagca aaagcatgag aatatatc 6960  
 aagagaaaaa tctaattatt ttataaaac aattatccca tgtaattat atgggaata 7020  
 tgataaatt gatgataact cgtcacatat gattccggca gttataaaaa aaatccatca 7080  
 tgcgaaaatt aataatgtcc cagagatcga aatttggggg gatggtaatt cgcccgctga 7140  
 gttctgtat gcagaagatt tagctgatct tattttttat gtatttctta aatagaatc 7200  
 catgcctaat atggttaaat ctggtttagg ttacgattat tcaattaatg actattata 7260  
 gataattgca gaagaaattg gttatactgg gagttttct catgatttaa caaaaccaac 7320  
 aggaatgaaa cggaagctag tagataattt attgcttaat aaaaattgtt ggccaagtca 7380  
 ctttgaactc agagatggca tcagaaagac ctataattat tacttgaga atcaaaata 7440  
 atgattacat acccatctgc tagtaatact tgggatgaat atgagtatgc agcaatcacg 7500  
 tcagtaattg actcaaaaat gttaccatg ggtaaaaagg ttgagttata tgagaaaatt 7560  
 tttgctgatt tgtttggtg caaatatgcc gtaatggta gctctggtc tacagetaat 7620  
 ctggttaatg ttgctgccct ttcttctact aataaaccaa aacttaaaag aggtgatgaa 7680  
 ataattgac ctgcaggttc atggtctacg acatattacc cctgcaca gtaggctta 7740  
 aaggtgaagt ttgctgatat caataaagaa actttaata ttgatatcga tagtttgaaa 7800  
 aatgctattt cagataaaaac aaagcaata ttgacagtaa atttattagg taactcata 7860  
 gatttttgcaa aaataaatga gataataaat aatagggata ttactctat agaagatac 7920  
 tgtgagtcga tgggcgcggt ctttcaaaat aagcaggcag gcacattcgg agttatgggt 7980  
 acccttagtt cttttactc tcatcatata gctacaatgg aagggggcgt cgtagtact 8040  
 gatgatgaag agctgtatca tgtattgttg tgccctcag ctcattggtg gacagaaat 8100  
 ttacaaaag agaatttggt tacaggcact aagagtgtg atattttcga agagtctgtt 8160  
 aagtttgggt taccaggata caatgttcgc ccacttgaaa tgagtgtgct tattgggata 8220  
 gagcaacta aaaaagtacc aggttttata tccaccagac gtccaatgc acaatattt 8280  
 gtgataaat taaagatca tccattcctt gatatacaaa aagaagtgtg tgaagtgcg 8340  
 tggtttggtt ttctctcgt tataaaggag ggagctgcta ttgagaggaa gagttagta 8400  
 aataactgta tctcagcagg cattgaatgc cgaccaattg ttactgggaa ttttctcaa 8460  
 aatgaacgtg ttttgagtta ttttgattac tctgtacatg atacggtagc aaagccgaa 8520  
 tatatagata agaattgggt tttgtcggga aaccaccaga tactttgtt taatgaaata 8580  
 gattattctc gaaaagtatt aaaaaacta acgaggcact ctatttcgaa tagagtgcct 8640  
 ttaagatggt attaacagtg aaaaaattt tagcctttag ctactctaa gactaccac 8700  
 cggttattga acagtttgtc aatccaattt gcatctctat tatcacacca ctaatactca 8760  
 accacctggg taagcaaaagc tatggtaatt ggaattttat aattactatt gtattttt 8820  
 ctcaagtta atgtggagga ttgtccgcat ggattgcaaa aatcattgca gaacagagaa 8880

ttcttagtga tttatcaaaa aaaaatgctt taegtcaaat ttctataat ttttcaattg 8940  
 ttattatcgc atttgcggtat ttgatttctt ttcttatatt aagtatttgt ttcttcgattg 9000  
 ttgcgaggaa taatttttca ttcttatctg cgattattat ttgtgggttt ttocaggaaag 9060  
 ttgataattt atttagttgt gcgctaaaag gttttgaaaa atttaagtga tcactgttttt 9120  
 ttgaagtaat tacaagagtg ctctgggctt ctatagtaat atatggcatt tacggaaatg 9180  
 cactcttata ttttacatgt tttagccttta ccattaaaag tatgctaaaa tatattcttg 9240  
 tatgtctgaa tattaccggt tgtttcatca atctcaattt taatagagttt gaggattgta 9300  
 atttgttaaa tgagtcaaaa tggatgtttc ttcaattaac tgggtggcgtc tcacttagtt 9360  
 tgtttgatag gctcgttaata ccattgattt tatctgtcag taaactggct tcttatgtcc 9420  
 ttgctcttca actagctcaa ttgatgttca ctctttctgc gtctgcaaat caaattattac 9480  
 tacciaatgt ttgtagaagt aaagcatcta acacatttcc ctctaattgt ttttttaaaa 9540  
 ttctgcttgt atcactaatt tctgttttgc ctgtcttgc gtatttcttt ttgggtcgtg 9600  
 atatatatc aatatggata aacctacat ttgcaactga aaattataaa ttaattgcaa 9660  
 tttttgctat aagttacatt ttattgtcaa tgatgacatc ttttcaattc ttgttattag 9720  
 gaattggtaa atctaaagctt gtgcaaat taaactgtgt tgcagggtcgc gcactgtcgt 9780  
 ctcaacggtt aatgcagctt cattatggcc ttatgcaat atctatggta aaaaataat 9840  
 atccgcttt tcaattttat tacctttatg tagcttttgt ctattttaat agagcgaaag 9900  
 atgtctattg atttactttt ttcaattact gaaatcgaa ttgttttttc ttgcaactatt 9960  
 tacatattta ctcaatgttt gttaatcggt aggatctatt tagataaaaat tattttaatt 10020  
 ctttttgct tgcctttttt tttagtaate attcaacttc ctgagcttaa tgtaaacggt 10080  
 ttgggtcgtt cttaaaagtt atcactgcct ttattgatgt ttcttatcgc ttttcaaaaa 10140  
 ccgaaattat gcttctgggt tattattgca ttgtgttttt tgaactctgc atttaatttt 10200  
 ttatatttaa agacatttga taagttagc tcatttctct ttactttttt tatatttctg 10260  
 ttttacttgt tttagatttga aattggtaat ttaccggttt ataaaaataa aaaaattttac 10320  
 gcgttgattt ttctctttat attaatagac ataatgcagt cattgttaat aaattatagg 10380  
 gggcagattt tatattccgt aatttgcac ctgatactgt tgtttaaagt taatttaaga 10440  
 aaaaagattc catacttttt tttaatgctg ccagttttat atgtaattat tatggcttat 10500  
 attgtgttta attatttcaa taaaggcgta actttttttg aaactacagc aagtaattat 10560  
 gaacgtacgg ggatgatata ttatttgggt tcacagctgt gtgattatat attccatgtt 10620  
 attggggacat taaatttctt aaataacggc ggacaatata agacgttata tggacttcca 10680  
 tcattaattc ctaatgacct tcattgattt ttattaccgt tctttataag tattgggtgtg 10740  
 ataggagcat tggttttatca ttctataatt ttgttttttt ttaggagaat acctttctta 10800  
 ttatatgaga gaaatgctcc tticattggt gtaagtgtgt tgttactgtt ccaagtgtgt 10860  
 ttaatttata cattaaaccc ttttgatgct tttaatcgat tgatttgcgg gcttacagtt 10920  
 gagggtgttt atggatttgc aaaaattaga taagtatacc tgtaattgaa atttagacgc 10980  
 tccacttgtt tcaataatca ttgcaactta taattctgaa ctgtatatag ctaagtgttt 11040  
 gcaatcggta actaatcaat cttataagaa tattgaaatc ataataatgt atggaggatc 11100  
 ttctgataaa acgcttgata ttgcaaaatc gtttaaagac gaccgaataa aaatagtttc 11160  
 agagaagaat cgttgaattt atgatgctgt gaataaagca gttgatttat ccattgtgtg 11220  
 ttgggttagca tttattgggt cagatgatgt ttactatcat acagatgcaa ttgcttctat 11280  
 gatgaagggg gttattggtat ctaattggcg ccctgtggtt tatggggagg cagcgacga 11340  
 aggtcccgat aggaacatat ctggattttc aggcagtgaa tgggtacaac taaacaggtt 11400  
 taagtttaat tattcaaaat gtaatttacc attgcccatt atgagcgcaa tatattctcg 11460  
 tgatttcttc agaaacgaac gttttgatat taaattaaaa attgttgcgt acgtgtgatt 11520  
 gtttctgtag ttgtttcatca aatggagtaa agagaagta ccttattttt taaatgacac 11580  
 gacccctatt tctagaatgg gatattgggt ggtttcgcact gatattttct ctcaagttaa 11640  
 aactacgcta gaaagtttca ttgtacgcaa aaagaataat atatcctgtt taaacatata 11700  
 gctgattctt agatatgcta aaatctcgtt gatggtagcg atcaaaaata tttttggcaa 11760

taatgtttat aaattaatgc ataacgggta tcattcccta aagaaaaatca agaataaaat 11820  
 atgaagattg tttatataat aaccggggctt acttggtggt gagccgaaca ccttatgagc 11880  
 cagttagcag accaaatggt tatacggcgg catgatgtta atattattgt tctaactggt 11940  
 atatcttgagg taaagcgcaac acaaaatatt aatatccatt atgttaatat ggataaaaaa 12000  
 tttagaagct ttttttagagc tttattttcaa gtaaaaaaaa taattgtcgc cttaaagcca 12060  
 gatataatca atagtcatat gtttcatgct aatattttta gtctgtttat taggatgtcg 12120  
 attccagcgg tgcacctgat atgtaccgca cacaacaaaa atgaaggtgg caatgcaagg 12180  
 atgttttgtt atcgactgag tgatttttta gcttctatta ctacaaatgt aagtaaaagag 12240  
 gctgttcaag agtttatagc aagaaaggct acacctaaaa ataaaatagt agagattccg 12300  
 aattttatta atacaaataa atttgatttt gatattaatg tcagaaagaa aacgcgagat 12360  
 gcttttaatt tgaaagacag tacagcagta ctgctcgag taggaagact tgttgagca 12420  
 aaagactatc cgaacttatt aaatgcaata aatcatttga ttccttcaaa aacatcaaat 12480  
 tgtaatgatt ttattttgct tattgtctggc gatggcgcat taagaaataa attattggat 12540  
 ttggtttgtc aattgaactt tgtggataaa gttttcttct tggggcaagg aagtgatatt 12600  
 aaagaattaa tgtgtgctgc agatcttttt gttttgagt ctgagtgagg aggttttgg 12660  
 ctggtgtgtg cagaagctat ggcgtgtgaa cgtcccggtg ttgctaccga tctctgtgga 12720  
 gtttaagaag tcgttggacc tcataatgat gttatccgt tcagtaatca tattctgttg 12780  
 gcagagaaaa tcgctgagac acttaaaata gatgataacg caagaaaaat aataggtatg 12840  
 aaaaatagag aatatattgt ttccaatttt tcaattaaaa cgaatgtgag tgagtgaggag 12900  
 cgcttatatt ttaaatattc caagcgtaat aataataatg attgaaataa taagtattga 12960  
 ctctggatgc aatagtttct ctatgctgtt tttttactgg ctccgtatatt ttacttatag 13020  
 ctggtatttt ttatatatca gtattaatct gtcctaact catctagact acattcaagc 13080  
 cgcgcatgct tcgctgctgc actacacctg acaggagatg gtaatgtcca agcaacagat 13140  
 cggcgtcgtc ggtatggcag tgatggggcg caacctggcg ctcaacatcg aaagccgcgg 13200  
 ttataccgtc tccattcttc accgctcccg cgagaaaaat gaagaagtgt ttgccagaa 13260  
 cccggataag aaactggctc ctattacac ggtgaaagag ttgcgtcagt ctcttgaaac 13320  
 cccacgtcgt atcctgttaa tggtaaaagc agggcgggga actgatgctg ctatcgatc 13380  
 cctgaagccg tatctggata aaggcgacat cattattgat ggtggcaaca cctctctcca 13440  
 ggacactatc cgtcgttaac gtgaactgtc cgcgggaaggc tttaacttca tcggtaccgg 13500  
 cgtgtccggc ggtgaagagg gcgcctgtaa aggcccatct atcgtccag tatggcagaa 13560  
 agaagcgtat gactgtgttg cgcctatcct gaccaagatt gctgcggttg ctgaagatgg 13620  
 cgaacctatg ataacttaca tcggtgtgta cggtcgggtg cactactgta agatgtgtga 13680  
 caacggtatg gaatatggcg atatgcagct gattgctgaa gctattcttc gctttaaagg 13740  
 cggccttaat ctgtctaacg aagagctggc aaccactttt accgagtggg atgaaggcga 13800  
 gctaagtagc taactgattg acatcaccaa agacatcttc accaaaaaag atgaagaggg 13860  
 taaaacctgt gttgatgtga tcttggaaga agctcgcaac aaaggcaccg gtaaatggac 13920  
 cagccagagc tctctggatc tgggtgaacc gctgtcgtg atcacgaat cgtatttcgc 13980  
 tcgctacatc tctctctgta aagaccagcg cattgcggca tctaaagtgc tgtctgtgct 14040  
 gcaggtctaa cgtgctgttg ataaagcaga gttcgttag aaagtccgtc gcgcgtgtga 14100  
 cctgggtaaa atcgtctctt atgcccaagg cttctctcaa ctgctgcgg cgtctgacga 14160  
 atacaactgg gatctgaact accggcgaat cgcgaagatc tccgcgcgg gctgcacat 14220  
 tcgtgcgag tctcgcaga aaattactga cgcgtatgct gaaaacaaag gcattgtctaa 14280  
 cctgtgtgct gctccgtact tcaaaaatat cgctgatgaa tatcagcaag cgtcgcgtga 14340  
 tgtagtggct tatgctgtgc agaaccgtat tccggtaccg acctctctg cagcggtagc 14400  
 ctactacgac agctaccgtt ctgcggtact gccggttaat ctgattcagg cacagcgtga 14460  
 ttacttcggt gcgcacacgt ataaacgcac tgataaagaa ggtgtgttcc acacgg 14516



&lt;211&gt; 1380

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 46

aacaatctc agtcttctct tagctctgct attgagcgct tgtcttcttg tctgcgtatt 60  
aacagcgcaa aagcagtgatc agcaggtcag gcgattgcta accgttttac ggcaaatatt 120  
aaaggtctga cccaggcttc ccgtaacgcg aatgatggta tttctgttgc gcagaccact 180  
gaaggtgcgc tgaatgaaat taacaacaac ctgcagcgta ttctgtgaact tctgttctag 240  
gcaactaacg gtactaactc tgacagcgat ctttcttcta tccaggctga aattactcaa 300  
cgtctggaag aaattgacgc tgtatctgag caaactcagt ttaacggcgt gaaagtccct 360  
gctgaaataa atgaaatgaa aattcagggt ggtgctaatt atggtgaaac catcactatc 420  
aatctggcaa aaattgatgc gaaaactctc ggcttgagc gttttaatat cgtatggcgc 480  
cagaaagcaa ccggcagtg cctgatttct aatttaaag cgacaggtac tgataattat 540  
caaattaacg gtactgataa ctatactgtt aatgtagata gtggagtagt acaggataaa 600  
gatggcaaac aagtttatgt gagtgctgcg gatggttcac ttacgaccag cagtataact 660  
caattcaaga ttgatgcaac taagcttgca gtgctgcta aagatttagc tcaaggtaat 720  
aagattgtct acgaaggtat cgaatttaca aataccggca ctggcgctat acctgccaca 780  
ggtaaatggg aattaacccg caatgttgat ggtaaggctg ttgaattcac tatttcgggg 840  
agtgctgata catcaggtac tagtgcaacc gttgcccta cgacagccct atacaaaaat 900  
agtcagggcg aattgatgcg acaaaaaggt gaaaataaag cagcgacact atctgatctt 960  
gatctgaacg ctgccagaa aacaggaagc acgttagtct ttaacggctg aactatcatg 1020  
gttagtgtag atggtaaaac gataacggag actgcttctg gtaacaataa agtcatgtat 1080  
ctgagcaaat cagaaggttg tagcccgatt ctggtaaaac aagatgcagc aaatcgcttg 1140  
caattctacca ccaacccgct cgaactatc gacaaagcat tggctaaagt tgacaactctg 1200  
cggtctgacc tcgggtcagc acaaaaccgt ttgcactctg ccatcaccac ccttgccaac 1260  
accgtaaaaa acctgtcttc tgcccgtagc cgtatcgaa atgctgacta cgcgaccgaa 1320  
gtgtctaaac tgtctctgct gcagatcctg caacaagcgg gtacctctgt tctggcacag 1380

&lt;210&gt; 47

&lt;211&gt; 1497

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 47

atggcacaa ctcattaatac caacagcctc tcgctgatca ctcaaaaataa tatcaacaag 60  
aaccagtcgt cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaaacagc 120  
gcgaaggatg acgcagcggg tcaggcgatt gctaaccggt tcacctataa catataaaggc 180  
ctgactcagg gggcccgtaa cyccaacgac ggtatctccg ttgcgcagac caccgaaggc 240  
gcgctgtccg aaatacaaca caacttacag cgtgtgcggt aactgacggt acaggccact 300  
accggtacta actctgagtc tgatctgtct tctatccagg acgaaataa atcccgctctg 360  
gatgaaatgt accgcgtatc tggtcagacc cagttcaacg gcgtgaacgt gctggcaaaa 420  
aatggctcca tgaaatcca ggttggcgca aatgataacc agactatcac tatcgatctg 480  
aagcagattg atgctaaaaa tcttggcctt gatggtttta gcgttaaaaa taacgataca 540  
gttaccacta gtgctccagt aactgctttt ggtgctacca ccacaaacaa tatataaact 600  
actggaatta ccctttctac ggaagcagcc actgatactg gcggaactaa cccagcttca 660  
attgaggggt tttatactga taatggtaat gattactatg cgaaaaatcac cggtgggtgat 720  
aacgatggga agtattacgc agtaacagtt gctaattgat gtacagtgc aatggcgact 780

ggagcaacg ccaatgcaac tgtaactgat gcaaatacta ctaaagctac aactatcact 840  
 tcaggcggtg cactgtgtca gattgataat actgcagggt ccgcaactgc caacctgtgt 900  
 gctgttagct tagtaaaact gcaggattcc aagggtaatg ataccgatac atatgcgctt 960  
 aaagatacaa atggcaatct ttacgctgcg gatgtgaatg aaactactgg tgcgtttctt 1020  
 gttaaaacta ttactatata tgactcttcc ggtgccgcca gttctccaac cgcggtcaaa 1080  
 ctgggcggaag atgatggcaa aacagaagtg gtcgatattg atggtaaaac atacgattct 1140  
 gccgatttaa atggcggtta tctgcaaaaca ggtttgactg ctgggtggta ggcctctgact 1200  
 gctgttgcaa atggtaaaac caccggtccg ctgaaagcgc tggacgatgc tatcgcatct 1260  
 tagacaaaat tccgttcttc cctcggtgcg gtgcaaaacc gtcctggattc ccggtgtacc 1320  
 aactgaaca acaccattac caacctgtct gaagcgcagt cccgtattca ggacgcccag 1380  
 tatgcgaccg aagtggtcaa tatgtcgaaa gcgcagatca tccagcagcg cggttaactcc 1440  
 gtgttggaac aagtaacca ggtaccgcag caggttctgt ctctgctgca ggggttaa 1497

&lt;210&gt; 48

&lt;211&gt; 1695

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 48

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaata tatcaacaag 60  
 aaccagtcgt cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120  
 gcgaaggatg acgcccaggg tcaggcgatt gctaacggtt ttactcttaa cattaaaggc 180  
 ctgactcagg tgcacgttaa cgccaacgac ggtatttctg ttgcgcagac caccgaaggc 240  
 gcgctgtctg aaatcaacaa caacttacag cgtattctgt aactgacggg tcagggtctt 300  
 accgggacta actctgatcc ggatctggac tccattcagg acgaaatcaa atcccgctgt 360  
 gcgaagaatt accgcgtatc cggtaaaacc cagttcaacg gtgtgaacgt actggcgaaa 420  
 gacggttcga tgaaaaattca ggttggtgcg aatgacggcc agactatcac tatgatctgt 480  
 aagaaaaatt actctgatac gctggggctg aatgggttta acgttaacgg caaagggtact 540  
 attgcgaaca aagcggcaac cattagtgtat ctggcgccga cggggcgcaa tgttactaac 600  
 tcaagcaata ttgtgttcac gacaaagttc aatgccttgg atgcagcgac tgcatttagc 660  
 aaactcaaaag atggtgatcc tgttgccgtt gctgctcaga aataacttta taacgcctcg 720  
 accaatgatt ttacgcagaa aaatacagta gcgacaggca ctgcaacgac agatcttggc 780  
 gctactctga aggcctgtgc tgggcagagt caatcaggta catataacct tgcaaagtgt 840  
 aaagttaact ttgatgttga tgcaagcggt aatatcacta ttggcgccga aaaggcttct 900  
 ttggttggtg gacgcgtgac tactaacgat cccaccgggt ccaactccag aacgatgtct 960  
 tccctgttta agcccgcca tgacaaaagt gccgctcaat cctcgattga ttttggcggt 1020  
 aaaaaatcac aatttgcgtg tggcaattct actaatggtg gcggcggttaa attcaagac 1080  
 acggtgtctt ctgacgcgtt ttggctcgag tttaaagcgg atagattcgc taataatgta 1140  
 aaaaacacct ttaacaactt tctctgttca ttactgcat gttccaaaaa tgggttatct 1200  
 ggctccgcgg catcgaatgc agcctacatt gatagcgaag gcgaactgac aactatgaa 1260  
 tctcacacaa caaattatcc cgtagacaaa gacacggggg ctgtaagtgt tacagggggg 1320  
 agcgggtacgg gtaaatcacg cgcaaacgtg ggtgctcagg cttatgtagg tgcagatggt 1380  
 aaattaacca cgaatactac tagtaccggc tctgcaacca aagatccact aaatgcgctg 1440  
 gatgaggcaa ttgcatccat cgacaaatcc cgttcttccc tgggggctat ccagaaccgt 1500  
 ctggattccg cagtcaccaa cctgaacaaac accactacca acctgtctga agcgcagtc 1560  
 cgtaattcagg acccgacta tgcgaccgaa gtgtccaaca tgcgaaagc gcagatcact 1620  
 cagcagcccg gtaactccgt gttggcaaaa gctaaccagg taaccgagca ggttctgtct 1680  
 ctgctgcagg gttaa 1695

- 37 -

&lt;210&gt; 49

&lt;211&gt; 1164

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 49

```

aacaagaacc agtctgcgct gtcgagttct atcgagcgct tgtctctctg cttgcgtatt 60
aacagcgcgga aggatgacgc cgccgggtcag gcgatttgcta accgcttttac ttctaaccatt 120
aaaggcctga ctcagcgtgc acgtaacgac aacgacggta tttctgttgc gcagaccacc 180
gaaggcgcgct tgtccgaaat taacaacaac ttacagcgct tgctgtgagct gactgttccag 240
gcgaccaccgc gtaactaacct tgagttctgac ctgtcttcta tccaggacga aatcaaatct 300
cgcttgggaag agattgatgc tgtttcaagt cagactcaat ttaacggcgt gaattgtttg 360
gctaaagatg ggaaaatgaa cattcaggtt ggggcaagtg atggacagac tatcattatt 420
gatctgaaaa agatcgattc atctacacta aacctctcca gttttgatgc tacaaccttg 480
ggcaccagtg ttaaagatgg ggccaccatc aataagcaag tggcagtaga tgctggcgac 540
tttaaagata aagcttcagg atcgtttagt accctaaaat tagttgagaa agacggtaag 600
tactatgtaa atgacactaa aagtagtaag tactacgatg ccgaagtaga tactagtaag 660
ggtgaaatac acttcaactc tacaatgtaa agtggaaacta ctctactgac agcgacggaa 720
gtaactactg ttggccgcga tgtaaaattg gatgctctcg cacttaaagc caaccaatcg 780
cttctcgtgt ataaagataa aagcggcaat gatgcttata tcattcagac caaagatgta 840
acaactaact aatcaacttt caatgccgct aatatcagtg atgctggtgt tttatctatt 900
ggtgcactca caaccgcgcc aagcaattta acagctgacc cgcttaaggg tcttgatgat 960
gcaattgcat ctgttgataa attccgctct tctctcggtg ccgttcagaa cgcctgggat 1020
tctgccattg ccaacctgaa caacaccact accaactctg ctgaagcgca gtcccgattt 1080
caggacgctg actatgcgac cgaagtgtcc aacatgtcga aagcgcatat tatccagcag 1140
gcgggtaact ccgtgctggc aaaa

```

&lt;210&gt; 50

&lt;211&gt; 1818

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 50

```

atggcacaag ctattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagctct cgctgtcgag ttctatcgag cgtctgtctt ctggctttgcg tattaacagc 120
gcgaaggatg acgcagcggg tcaggcgatt gctaaccggt tcacctctaa cattaaaggc 180
ctgactcagg ctgcacgtaa cgtcaacgat ggtatctctc tggcgacagc cactgaaggc 240
gcactgtctg agatttaaca caacttaaca cgtgtgcgtg agttgactgt acaggcgacc 300
accggtacta actctgatcc tgacctggct tctattcagg acgaaatcaa atcccgtttg 360
tctgaaattg acccggtatc cgggcagacc cagtccaacg cgctgaacgt attgtctaaa 420
gatggctccc tgaataattc ggttggcgca aatgatggtc agactatctc tatcgacctg 480
aagaaaattg actctgatac tctgggtttg aatggtttca acgttaatgg ttctgggtacc 540
attgcaaaac aagcgccacc aatcagtgac ttgactgctc agaaagccgt tgacaacggt 600
aatggctact ataaagttac aactagcaac gctgcactta ctgcatctca ggcatthaag 660
aagctgagtg atggcgatac tgtagatatt gcaacctatg ctggttggtac aagttcaaca 720
gttagttata aatcgcacgc agatgcaggt aacttcagtt ataacaatac tgcacaacaa 780
acaagtgcgt cggctggaaac tctggcagat actctctccc cggcagctgg ccagactaaa 840

```

accggtactt acaaggctgc tactggtgat gttacttta atgttgacgc aactggtaat 900  
 ctgacaattg gcggacagca agcctacctg actactgatg gtaaccttac aacaacaac 960  
 tccggtggtg cggtactatg aactcttaaa gagctgttta ctcttgctgg cgatggtaaa 1020  
 tctctgggga acggcggtac tgcctaccgtt actctggata atactacgta taatttcaaa 1080  
 gctgctcgga acgttactga tgggtctggt gtcacgctg ctgctggtgt aacttataca 1140  
 gccactgttt ctaaaagtgt cattctggca caactgcaat ctgcaagtc ggcagcagca 1200  
 accgctaccg acgggtgatac tgcgcacacg atcaactata aactctggtg catgatcggt 1260  
 tccgctacct ttaccaatgg taaaggtaact gccgatggtg tgacttctgg tacaactcca 1320  
 tcgttagcta caggtgctaa agctgtatat gttgatggca acaatgaact gacttccact 1380  
 gcattctacg atacgactta ctctgtcaac gcagatacag gcgcagtaaa agtggatca 1440  
 ggtactggta ctggtaaatt tgaagctggt gctgggtcgg atgcttatgt aagcaaatgt 1500  
 ggcaaatata cgacagaaac caccagtgcg ggcaactgca ccaaatgctc ttgggtgccc 1560  
 ctggatgctg ctatcagctc catcgacaaa ttccgttctc cctgggtgct tatccagaac 1620  
 cgtctggatt ccgacgtcac caacctgaac aacaccacta ctaacctgtc tgaagcgcag 1680  
 tcccgatttc aggaacgcga ctatcgacac gaagtgtcca atatgtcgaa agcgcagatc 1740  
 atccagcagg ccggttaact tgtgttgcca aaagctaacc aggtaccgca gcaggttctg 1800  
 tctctgctgc aggggttaa 1818

&lt;210&gt; 51

&lt;211&gt; 1344

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 51

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaata tatcaacaag 60  
 aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcc tattaacagc 120  
 gcgaaggatg acgcgcagg tcaggcgatt gctaacggtt tacttctaaa ctataaaggc 180  
 ctgactcagg ctgcacgtaa cgccaacgac ggtattttctg ttgcacagac cactgaaggc 240  
 gcgctgtccg aaatcaacaa caacttacag cgtattctgt aactgacggt tcaggttctt 300  
 accgggacta actctgatcc ggaatctggac tccattcagg acgaaatcaa atccgtctc 360  
 gacgaaattg accgcgttcc cgttcagacc cagttcaacg gcgtgaacgt gctggcgaaa 420  
 gacggttcga tgaagattca ggttggcggc aatgacgggc agaccatctc tatcgatttg 480  
 cagaaaaattg attcttcaac gctgggattg aaaggtttct cggtatcagg gaacgcatta 540  
 aaagttagcg atgcgataac tacagttctc ggtgctaatt ctggcgatgc cccggttacg 600  
 gttaaatgtg gtgcgaacga taccgtgctt gccgcaatg ctaaaacatt gggaataagt 660  
 gatacatcag gcttctccct acataacgta caaagcgcgg atggtaaagc gcaggaac 720  
 tatgttgttc aatctggtaa tgacttctat tcggcttccg ttaatgctgg tggcggtgtt 780  
 acgcttaata ccaccaatgt tactttcact gatcctgcga ccgctgttac cacagaacaa 840  
 cagcagctgc agcctatcaa ggctcagcag aatagtgctg gcgcggctgt tgctatggtt 900  
 actattcaag gcgaagatta ccttctggtt gcagacggtg aggatgcaat tgaacacggt 960  
 ggtgacgtct gacaaaatga agacacaaaa atccaactta ccgatgaact cgatgttgat 1020  
 ggttctgttaa aaacagcgcg aacagcaaca tttcttggtg ctgcaaccaa cgatccgctg 1080  
 gcacttttag acaaacgtat ctcgcaagtt gatactttcc gctctctcct cggtgccgta 1140  
 caaacaccgtc tggattctgc ggtcaccacac ctgaataaca ccaccaccaa cctgtctgaa 1200  
 gcgcagtcct gttattcagga cgcgactat gcgacccaag tgtccaacat gtcgaaagcg 1260  
 cagatctacc agcagcgcgg taactctgtg ctgtctctaa ctaaccaggt accgcagcag 1320  
 gttctgtctc tgcgtcaggg ttaa 1344

&lt;210&gt; 52

- 39 -

&lt;211&gt; 2599

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 52

cttctcttag ctctgctatt gagcgtctgt cttctggtct gcgtattaac agcgcaaaag 60  
 acgatgcagc aggtcaggcg attgctaacc gttttacggc aaatattaaa ggtctgaccc 120  
 aggcttcccg taacgcgaat gatggtattt ctgttcgcga gaccactgaa ggtgcgctga 180  
 atgaaattaa caacaacctg cagcgtatct gtgaactttc tgttcaggca actaacggta 240  
 ctaactctga cagcgatctt tcttctatcc aggcgtgaaat tactcaacgt ctggaagaaa 300  
 ttgacgctgt atctgagcaa actcagttta accgctgtaa agtccttctg gaaaaaatg 360  
 aaatgaaat tcaggttggt gctaattgat gtgaaacat tgacctgccc ccacgattag 420  
 atacaacact cagtttagta cgtcggaatc ttcattctca gaatgacct ttctccagcc 480  
 cgctgcaaat tcagacgggt tctgataatt cagcgtggag tgcggggcgc attcgttata 540  
 atcttgcgc cagtcattaa taattttcct ggcatagaac atactcgctga accagtgctc 600  
 attcaaacat tcatcgcgaa atcgtccgtt aaagctctca ataaatccgt tctgcgttgg 660  
 cttgcccgcc tggattaagc gcaactcaac accatgctca aaggcccat gatccagtgc 720  
 accgcaagtg aactccggcc cctggtcagt tcttatcgtc gccgatagc ctcgaaacag 780  
 tgcaatgctg tccagaatac cggtgacctg aacgcctgaa atccccaaag caacagtgc 840  
 cgtcaggcat tcttttgta aatcatcgac gcaggttaaga cacttgatcc tgcgacggt 900  
 ggaaagtgcg tccatgacga aatccatcga ccaggtcaga ttggcgccgc ccggaacggag 960  
 cagcggcaga cgttctgttg ccagcccttt acgacgtctt ctgcgtttta cggccaggcc 1020  
 actgaggtga taaagccgtt acacgcgctt atgattaaac tgaagccctt cagcgcgca 1080  
 caactgccaa atacgcaggt agccaaaacg cctgcgctcc agtgccagct cagtgatgcg 1140  
 cctgataaaa tgcgcacatg cagccggacg gtgagctca tagcggcagg tcgacagggg 1200  
 taaacctgta agctgcaggt cagcagcttg cgacagaccg gtgcatacac acatcaacat 1260  
 cagcgcttcc cgcttctggt ctgtcgtcag tactttcgcc caagagccac ctgaagcgcc 1320  
 tctttatcca gcatggcttc ggcaagcagc ttcttgagtc tgggtgtctc ttctcaagc 1380  
 gacttcagcg ccttaacctt aggcacctcc ataccgccat acttcttagc ccaggtgtta 1440  
 aacgtggcat cggaaatggc atgcttgcgg cagagttcac gggcgggatc ccagcttcg 1500  
 gcttcgcgga gaatactgat gatctgttcg tcggaaaaac gcttctcat ggggatgtcc 1560  
 tcatgtggct tatgaagaca ttactaacat cgggggtgtac taatcaacgg ggagcaggtc 1620  
 accatcacta tcaatctggc aaaaattgat gcgaaaactc tcggccttga cggttttaat 1680  
 atcgatggcg cgcagaaagc aaccggcaggt gacctgatt ctaaatttaa agcgacaggt 1740  
 actgataatt atcaaattaa cggtaactgat aactatactg ttaattaga tagtggagta 1800  
 gtacaggata aagatggcaa acaagtttat gtgagtgctg cggatggttc acttcagacc 1860  
 agcagtgata ctcaattcaa gattgatgca actaagcttg cagtggctgc taaagattta 1920  
 gctcaaggata ataagattgt ctacgaaggt atcgaaattt caaataccgg cactggcgct 1980  
 atacctgcca caggtaatgt taaattaacc gccaatgttg atggttaaggc tgttgaattc 2040  
 actatttcgg gtagtgctga tacatcaggt actagtgcga cegtgcgcc tacgacagcc 2100  
 ctatacaaaa atagtgcagg gcaattgact gcaacaaaag ttgaaaaata agcagcgaca 2160  
 ctatctgatc ttgatctgaa cgctgcgaag aaaaacaggaa gcacgttagt tgttaacggt 2220  
 gcaacttacg atgttagtgc agatggtaaa acgataacgg agactgcttc tggtaacaat 2280  
 aaagtcatgt atctgagcaa atcagaaggt ggtagccgga tcttggtaaa cgaagtgcga 2340  
 gcaaaatcgt tgcaatctac caccaccccg ctgcaaaact tcgacaaagg attggtctaa 2400  
 gttgacaatc tgcgttctga cctcggtgca gtacaaaacc gtttcgaact tgccatcacc 2460  
 aaccttgga acaccgtaaa caacctgtct tctgcctga gccgatcga agatgctgac 2520

tacgcgaccg aagtgtctaa catgtctcgt gcgcagatcc tgcaacaagc gggtagctct 2580  
gtttcggcac aggctaacc 2599

<210> 53

<211> 1245

<212> DNA

<213> *Escherichia coli*

<400> 53

aacaaaaacc agtctgcgct gtgcacttct atcgagcgcc tctcttctgg tctgcgcatt 60  
aacacgcgcta aagatgacgc tgcgggcccag gcgattgcta accgcttcac ttctaacatc 120  
aaaggtctga ctacaggccgc acgtaacgcc aacgacggta tctctctggc gcagaccact 180  
gaagggcgac tgtctgaat caacaacaac ttgcagcgtg ttctgtaact gaccgttcag 240  
gccactaccg gtaactaact tgattctgac ctgtcttcaa tccaggacga aatcaaatcc 300  
cgtctcgtag aaattgacgc cgtatccggc cagactcagt tcaacggcgt gaacgtactg 360  
gcaaaagacg gctcgtgtaa aattcaggtc ggtgcaaatg atggtcagac aatcagcatt 420  
gatttcaga agattgattc ttctacttta ggggttaaat gttttctgt ttccaaaaat 480  
gcagtatctg ttggtgatgc tattactcaa ttgcttggcg agacggcgcg cgatgcacca 540  
gtaacatca agtttgatga ttcatgaaaa actgatttaa aactgaccga tgcctcaggg 600  
ttaagtctgc ataactcaa agatgaaaat ggttaattaa ctaaccagta tgggtacag 660  
aatggcgcaa aatcttacgc tgctacagtc gctgccaatg gtaattgtac cgtgaacaaa 720  
gcaaatgtaa cctacagcga tgcgcgaac ggtattgata ccgcaacgca gtcaggccag 780  
ttagtccagg ttggtgcaga ttctaccggc acgccaagaag cattcgtgtc tgtccaagg 840  
aaaagcttgg gcattgatga cgcgcgcttg aagaataaca ctggtgtatgc taccgctact 900  
ccaccgggaa catctgggac aacagttgtc gcagcgtcaa ttcatctgag tacgggcaaa 960  
aactctgtag acgtgtatgt aacggcttcc actgaattca caggtgtctc aaaccaagat 1020  
ccactgactc tgctggacaa agctatcgca tctgtgata aatctcgttc ttcttgggg 1080  
gcggtacaga accgtctgag ctccgctgta accaacctga acaaacaccac caccacctg 1140  
tctgaagcgc agtcccgatc tcaggagccc gactatcgca ccgaagtgtc caacatgtcg 1200  
aaagcgaga ttatccagca ggcaggtaac tccgtctgtg ccaa 1245

<210> 54

<211> 1212

<212> DNA

<213> *Escherichia coli*

<400> 54

aacaaaaacc agtctgcgct gtgcacttct atcgagcgcc tctcttctgg cctgcgtatt 60  
aacagtgcca aagatgacgc tgcgggtcag gcgtagctca accgtttcac ctctaactatt 120  
aaaggcctga ctacaggctgc gcgtaacgcc aacgacggta tttctctggc gcagaccaca 180  
gaaggtgcgt tgtctgaat caacaacaac ttgcaacgtg tgctgtagtt gaccgttcag 240  
gcgacgacgc gtaactaact tgattctgac ctgtcatcta ttacgacga aatcaaatcc 300  
cgtctggatg agattgacgc gttttccggc cagacccagt tcaacggcgt gaattgactg 360  
gcaaaagacg gttcgtatga gattcagggt ggcgcgaatg atggccagac tattagcatt 420  
gatttcaga aaattgactc ttctacatta ggggtgaatg gttctcctgt ttctgtctca 480  
tcaactaacg ttggtgatcc aattactcaa attacaggag ccgctgggac aaaacctgtt 540  
ggtgttgatt tcaactgtgt tgcaaaagat ctgactactg cgcagggtaa aactgtcgat 600  
gtttccagcc tgacgttaca caacacctg ctgcgcaaa gggctgccac cgcacagttc 660

gtcgttcaat ccggttagta tttctactcc gcgtccattg accatgcaag tgggtgaagt 720  
 acgttgaata aagccgatgt cgaatacaaa gacaccgata atggactaac gactgcagct 780  
 actcagaaga atcagctgat taaagttgcc gctgactctg accgcccggc tgcgggatat 840  
 gtaacattcc agggtaaaaa ctacgctaca accgctccag cggcgcttaa tgatgacact 900  
 accgcaacag ccacagcgaa caaagtgttt gttgaattat ctacagcaac tccgactgog 960  
 cagttctcag gggctctctt tgctgatcca ctggcacttt tagacaaaag cattgcacag 1020  
 gttgataact tccgctctct cctcggtgcc gttcaaaacc gtctggactc tgcggtaacc 1080  
 aactgaacca acacacacac caacctgtct gaagcgcagt cccgtattca ggacgcccag 1140  
 tatgcgacgc aagtgtctaa catgtcgaaa gcgcagatca tccagcagggc gggtaactct 1200  
 gtgctgtcta aa 1212

&lt;210&gt; 55

&lt;211&gt; 1758

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 55

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaata tatcaacaag 60  
 aaccagtcct cgtctgcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120  
 gcgaaggatg acgcccgagg tcaggcgatt gctaaccggt ttacttctaa cattaagaagg 180  
 ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcacagac cactgaaggc 240  
 gcgctgtccg aaatacaaaa caacttacag cgtatccgtg agctgacggt tcaggctctct 300  
 accgggacta actctgattc ggatctggac tccattcagg acgaatacaa atcccgcttc 360  
 gacgaaattg accgcgtatc cggtcagacc cagttcaacg cgggtgaact actggcaaaa 420  
 gacggttcga tgaataatca ggttggtgag aatgacgggt aaactatcac tatcgacctg 480  
 aagaaaaatc attctgatac tctgggtctg aatggtttta acgtaaatgg taaaggtact 540  
 attaccaaca aagctgcacac ggtaagtgat ttaacttctg ctggcgcgaa gtttaaacac 600  
 acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgcctgattc 660  
 gataaattag ggaatggcga taaagtcacc gttggcgagg tagattatac ttacaacgct 720  
 aaatctgggt attttactac caccaaatct actgctggta cgggtgtaga cgccgcggcg 780  
 caggctactg attcagctaa aaaacgtgat cggcttagct ccaaccttca tgcctgatgt 840  
 ggtaaatctg ttaatgggtc ttacaccaca aaagatggta ctgttctctt cgaaacggat 900  
 tcagcaggta atatcaccat cggtggaagc caggcatacg tagacgatgc aggcaacttg 960  
 acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgtgctg taaagccgcg 1020  
 agcgaaggta gtgacgggtc ctctctgaca ttcaatggca ctgaatatat tatcgcaaaa 1080  
 gcaactcctg cgacaacctc tccagtatgt ccgttaatcc ctgggtggat tactatcagc 1140  
 gctacagtga gtaaatgatg agtattgagc gaaaccaaa cggctgcgcg gacattctca 1200  
 attaccttta attccgggtg actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260  
 gatgcgcgga atctctatgt ggatgataaa ggtggattta ctacgctgc cgactataca 1320  
 gtctcttaca cgttatacaa ggataacggc tctgtgactg ttgcccgggt tgcctcagcg 1380  
 actgatacca ataagatta tgctccagca attgggtact ctgtaaatgt gaactccgcg 1440  
 gtaaaaaata ctactgagac taccagtgct ggttctgcga cgaccaacct gcttgctgcc 1500  
 ctggacgacg ctatcagctc catcgacaaa ttccgttctt cctcggtgct tatccagaac 1560  
 cgtctggatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc tgaagcgcag 1620  
 tcccgtaata aggaagccga ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatt 1680  
 attccagcag ccggttaact cgtgctggca aaagccaacc aggtaccgca gcaggttctg 1740  
 tctctgtctg agggttaa 1758

&lt;210&gt; 56

&lt;211&gt; 14024

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 56

gtaaccaagc gccgtacgtg cataaaatttt aatgctttac aaaactatta gcattaaaaa 60  
tatataagaa attctcaaat gaacaaagaa accgtttcaa taattatgcc cgtttacaat 120  
ggggccaaaa ctataatctc atcagtagaa tcaattatcc atcaatctta tcaagatttt 180  
gttttgtata tcatcgaca ttgtagcacc gatgatacat tttcattaat caacagtcga 240  
tacaaaaaca atcagaaat aagaatattg cgtaacaaga caaatttagg tgttgcagaa 300  
agtcgaaatt atggaataga aatggccacg gggaaatata tttctttttg tgatcgcat 360  
gatttgtggc acgagaaaaa attagagcgt caaatcgaag tgttaaataa tgaatgtgta 420  
gatgtgggtat gttctaatta ttatgttata gataacaata gaaatttgt tggcgaaagt 480  
aatgctcctc atgtgataaa ttatagaaaa atgctcatga aaaactacat aggggaattg 540  
acaggaatct ataattgcaa caaattgggt aagttttatc aaaaaaagat tggtcacgag 600  
gatttattga tgtggctgga aataattaat aaacaaatg gtgctatttg tattcaagat 660  
aatctggcgt attacatgcg ttcaataat tcaactatcg gtaataaaat taaagctgca 720  
aaatggacat gggatgata tagagaacat ttacatttgt cctttccaaa aactatatat 780  
tattttttat tatatgcttc aaatggagtc atgaaaaaaa taacacatct actattaagg 840  
agaaaggaga ctaaaaagtg aagtcagcgg ctaagttgat ttttttattc ctatttacac 900  
tttatagctc ccagttgtat ggggttatca tagatgatcg tatacaaaat ttgtatacca 960  
aggattatac tagtattata attatatttc agattttttt ttttttctaa 1020  
cgattataaa tgaagaagaa cagcagaaaa aatttatcgt gaactgggag ctaaaagtta 1080  
tactcgtttt ccttttttgt actatagaaa ttgctgctgt agttttattt cttaaagaag 1140  
gtattccatt atttgatgat gatccagggg gggctaaact tagaatagct gaaggtaatg 1200  
gactttacat tagatatatt aagtattttg gtaatatagt tgtgtttgca ttaattattc 1260  
tttatgatga gcataaattc aaacagagga ccatcatatt tgtatatttt acaacgattg 1320  
ctttatttgg ttatcgttct gaattgggtg tgctcattct tcaatatata ttgatacca 1380  
atatcctgtc aaaggataac cgtaatccta aataaaaaag aataataggg tattttttat 1440  
tggtaggggt tgtatgctcg ttgtttttatc taagttttagg acaagacgga gaacaaatg 1500  
actcatataa taatatgtta aggataatta ataggtttaa aatagagcaa gttgaaggtg 1560  
ttccatagtg tgtttctgaa tctattaaga acgatttctt tccgacacca gaggtagaaa 1620  
aggaattaaa agcaataata aatagaatac aggggaataa gcatcaagac ttattttatg 1680  
ggaacgggtt acataaacaa gtatttggag acatgggagc aaatttttta tcaagtacta 1740  
cgtagggagc agaaactgta gttttttttg gttttctctg tgtattcaat tcccttttag 1800  
ggatatatat acctttttat cttttaaaga gaatgaaaaa aaccatagc tcgataaatt 1860  
gcgcattcta ttcatatatc attatgattt tattgcaata cttagtggct ggggaatgat 1920  
cgccctttct ttttggctct ttctctccg tattgataat gtgtactcct cgtactttat 1980  
tgcatgatgc gttaaagaga ttatcacgaa atgaaaatat cagtataaac tgtgacttat 2040  
aataatgctg aagggttaga aaaaacttta agtagtttat caattttaaa aataaaacct 2100  
tttagagatta ttatagttga tggcggtctc acagatggaa cgaatcggtg cattagtaga 2160  
tttactagta tgaattattc acatgtttat gaaaaagatg aagggatata tgatgcgatg 2220  
aataagggcc gaatgttggc caaaggcgac ttaatacat atttaaacgc cggcgatagc 2280  
gtaattggag atataataa aaatatcaaa gagccatgtt tgattaaagt tggccttttc 2340  
gaaaatgata aactctcggg attttcttct ataaccatt caaatacagg gtattgtcat 2400  
caaggggtga ttttccaaa gaatcatcca gaatatgatc taaggtataa aatatgtgct 2460  
gattataagc ttatcaaga ggtgtttctc gaagggttaa gctctctac ttgtattact 2520



tcgggttatg taaaatatga tatgggggga gtatcttcaa aaaaaagaat ttttaagagat 2580  
 aaagagcttg ccaaaatgat gtttgaaaaa aataaaaaaa accattattaa gtttattcca 2640  
 atttcaataa tcaaaatatt atccctgaa cgtttaagaa gagtattgag gaaaatgcaa 2700  
 tatatttgct taactttatt cttcatgaag aatagtacac catatgataa tgaataaaat 2760  
 caaaaaaata cttaaatatt gcaactttaa aaaatatgat acatcaagtg ctttaggtag 2820  
 agaacaggaa aggtacagga ttatatcctt gtcgtttatt tcaagtttga ttagtaaaat 2880  
 actctcacta cttctcttta tattaactgt aagtttaact ttaccttatt taggacaaga 2940  
 gagattttgt gtttgatga ctattaccag tcttgggtgt gctctgactt taatagactt 3000  
 aggtatagga aatgcattaa caaacaggat cgcacattca tttgcgtgtg gcaaaaattt 3060  
 aaagatgagt cggcaaatga gtgggtgggt cactttgtgt gctggattat cgtttgtcat 3120  
 aactgcaata tcttatatta cttctggcat gattgattgg caactagtaa taataaggtat 3180  
 aaacgagaat gtgtatgcag agttacaaca ctcaattaaa gtctttgtaa tcatatttgg 3240  
 acttggaaat tattcaaatg gtgtgcaaaa agtttatatt ggaatacaaa aagcctatat 3300  
 aagtaattat ttttaatgca tatttatatt ttatctattt attactctag taatatctgc 3360  
 gaaactacat ggggactac cagttttaat gtgcagcact cttgggtatt aatacatatc 3420  
 ggggaatcct ttacaatata atcttattat aaagcgatta ataaagttaa caaagtttaa 3480  
 catatcgtct aaaaagaaag ctccatattt gatataaaac ggttttttct tttttatttt 3540  
 acagttagcg actctggcaa catggagtgg tgataacttt ataatactta taacattggg 3600  
 tgttacttat gttgctgttt ttagcattac acagagatta tttcaaatat ctacggtccc 3660  
 tcttacgatt taaacatcc cgtttatggc tgcttatgca gatgtcatg cagcgaatga 3720  
 tactcaattt ataaaagaag cgtcgaagac atcattgaaa atagtgggta tttcatcatt 3780  
 cttattggcc ttcatattag tagtggtcgg tagtgaagtc gtttaatttt ggacagaaag 3840  
 aaagattcag gtacctcgaa cattcataat agctttatgt ttatggtctg tttattgatg 3900  
 tttttcgaat acttttgcaa gctttttaa atagttgaac taagttaaac aacaaatgct 3960  
 tggctgttga acattgatat tgatcgcaat tccagcaaaa tacatcatag tttagccatt 4020  
 tgggttaact gttatgtgt actgcttcat ttttatatat attgtaaat accatttatg 4080  
 gtataaatgt agttttaaaa aacatatcga tagacagtta aataaagag gtagaaatg 4140  
 aaatatatac cagtttacc aacgtcattg acaggaaaag aaaaagaata tgtaaatgaa 4200  
 tgtctggagt caacgtggat ttcatcaaaa ggaactata ttcagaagtt tgaaaaataa 4260  
 tttgcggaac aaacacatgt gcaatatgca actactgtaa gtaatggaac ggtgtccttt 4320  
 catttagcct tgttagcgtt aggtatatcg gaaggagatg aagttattgt tccaacactg 4380  
 acatatatag catcagttaa tgctataaaa tacacaggag ccaccccat tttcgttgat 4440  
 tcagataatg aaacttgcca aatgtctgtt agtgacatag aacaaaaaat cactaatana 4500  
 actaaagcta ttatgtgtgt ccatttatac ggacatccat gtgatatgga acaaatgtga 4560  
 gaactggcca aaagtatgaa tttgtttgta attgaagatt ggcgtgaagc ctttggttct 4620  
 aaatataaag gtaaaatgt gggaaacatt ggagataatt ctacttttag ctttttga 4680  
 aataaaacta ttactacagg tgaaggtgga atggttgta cgaatgacaa aacactttat 4740  
 gaccgttgtt tacattttaa aggccaaaga ttagctgtac ataggcaata ttggcatgac 4800  
 gtatataggt acaattatag gatgacaaat atctgcgtg ctataggatt agccagttga 4860  
 gaacaaagctg atgattttat atcacgaaaa cgtgaaattg ctgatattta taaaaaaat 4920  
 atcaacagtc ttgtacaagt ccacaaggaa agtaaaagat tttttcacac ttttggtatg 4980  
 gtctcaattg taactaggac cgcagaggaa agagaggaaat taaggaaata ccttgcatg 5040  
 aaactcatcg aaacaaggcc agttttttac cctgtccaca cgtatgcaat gtactcgga 5100  
 aaatatcaaa agcacccat agctgaggat cttggttggc gtggaattaa ttactactgt 5160  
 tctcccgagc tatcgaaatg gcaagttatt tatatttctg aatcatataa cgaattttat 5220  
 agtgataaat agcctaaaa attgtaaagg tcatctatga aaattgcgtt gaattcagat 5280  
 ggaattttac agtggggcgg tggaaattgat tttatataat atattctgtc aatattagaa 5340  
 acgaacacag aaatatgtat cgaattcttt ttaccgagaa atgatataca tttcttata 5400

agagaaaaag catttccttt taaaagtata ttaaaagcaa ttttaagag ggaaggcct 5460  
 cgatggattt cattaaatag atttaatgag caatactata gagatgcctt tacacaaaat 5520  
 aatatagaga cgaatctttac ctttattaaa agtaagagct ctgcctttta ttcatatttt 5580  
 gatagtgcgc attgtgatgt tatttctcct tgcattgcgtg ttctctcggg aaatttgaaat 5640  
 aaaaaagcat ggattgggta tatttatgac tttaacact gttactatcc ttcatttttt 5700  
 agtaagcgag aaatagatca aaggaatgtg ttttttaaat tgatgctcaa tgcgctcaac 5760  
 aatattattg ttaattgcaca ttcatgtatt accgatgcaa ataaatagt ttgggaattat 5820  
 tctgcaaaac tacatttctt tccatttagt ccatgccttc aattaaaatg gttcgtgat 5880  
 tactctggta atattgcaa atataatatt gacaaggatt attttataat ttgcaatcaa 5940  
 ttttggaaac ataaagatca tgcaactgct tttagggcat taaaattta tactgaatat 6000  
 aatcctgatg tttatttagt atgcacggga gctactcaag attatcgatt ccttggtat 6060  
 tttaatgaat tgatgggttt ggcaaaaaag ctcggaattg aatcgaaaat taagatatta 6120  
 ggcatataac ctaaacttga acaaatgaa ttaatcaaaa attgcattgc tgtaatacaa 6180  
 cccaacttat ttgaaggcgg gcttggaggg ggggttaacat ttgacgtat tgcatagggg 6240  
 aaaaaagtta tactatctga catagatgtc aataaagaag ttaattgcgg tgatgtatat 6300  
 ttcttctcgg caaaaacca ttattcatta aatgacgga tggtaaaagc tgatgaatct 6360  
 aaaatttttt atgaactcac aactctgata gaattgggtc tcaaaagacg caatgcgtgt 6420  
 gcagattttc ttttagatgt tgtgaaacaa gaaattgaat cccgatctta atattattca 6480  
 gaggtatata atgactaaag tcgctcttat tgcaggtgta actggacaag atggatctta 6540  
 cttagctgag ttttctgtg ataaagggta tgaagttcat tgcataaac gccgacgtc 6600  
 atcttttaat acagaacgca tagaccatat ttatcaagat ccacatgggt ctaacccaaa 6660  
 ttttccctg cactatggag atctgactga ttactataac ctactagaa ttctaagaa 6720  
 ggtacgcga gctgaagtat ataatttagc tgcattagat cactagcagc taatttttga 6780  
 gtctccagaa tatacagcgc atgtcgtatc aattgggtaca ttactgttac tgggaagcaat 6840  
 tcgcttttta ggattgaaaa acaaaacgcg ttctatcaa gcttcaacct cagaattata 6900  
 tggacttgtt caggaaaatcc ctcaaaaaag atccacctt ttattctc ttctcctta 6960  
 tgcagttgca aaactttacg catattggat cccggttaat tatcgagat catatggtat 7020  
 tatgcatgt aatgttatat tgttcaatca tgaattctca cgcgtggag aaacgtttgt 7080  
 aacaaggaaa attactcgag gacttgcaaa tattgcacaa ggcttggaaat catgtttga 7140  
 tttaggaat atggattcgt tacgagattg gggacatgca aaagattatg ttagaatgca 7200  
 atggttgatg ttacaacagg agcaaccgca agattttgtg attgcaacag gactccaata 7260  
 ctacgtccgt cagtttgcg aaatggcagc agcacaactt ggtattaaga tgagctttgt 7320  
 tggtaaaagga atcgaaagaaa aaggcattgt agattcggtt gaaggacagg atgctccagg 7380  
 tgtgaaacca ggtgatgtca ttgttgcgtg tgactcctgt tatttccgac cagctgaaat 7440  
 tgatactttg cttggagatc cgagcaaaagc taacttcaaa cttgggttga gaccagaaat 7500  
 tactcttgcg gaaatgattt ctgaaatggt tgcaaaagat ctgaagcgg ctaaaaaaca 7560  
 ttctctttta aaatcgcatg gttttctgtg aagcttagct ctggaatgat gatgaataag 7620  
 caacgtattt ttattgctgg tcaccaagga atgggtggat cagctattac ccgacgcctc 7680  
 aacaacacgt atgatgttga tttggtttta cgtactcagg atgaattgaa ctgttggat 7740  
 agtagcgtg ttttggattt tttttcttca cagaaaatcg accaggttta tttggcagca 7800  
 gcaaaagtcg gaggattttt agctaacagt tcttactcgt ccgattttat atatgagaa 7860  
 ataatgatag aggcgaatgt cattcatgct gccacaaaa ataatgtaaa taactgcctt 7920  
 ttctcgtggt cgtcgtgtat ttactctaag tttagcacac aaccgattat ggaagacgaa 7980  
 ttattacaag ggaacttga gccaaacaaat gaaccttatg ctatcgcmaa aattgcaggt 8040  
 attaaattat gtgaaattta taacogtcag ttgtggcggt attaccgttc agtaatgcc 8100  
 accaatcttt atggtccaaa tgacaatttt catccaagta attctcatgt gattccggcg 8160  
 cttttgcgcg gctttcatga tgcgtgtgaa aacaattctc cgaatgtgt tgtttgggga 8220  
 agtggtaact caaagcgtga attcttcatg ttagatgata tggcttctgc aagcatttat 8280

gtcatggaga tgcacatcga tatatggcaa aaaaatacta aagtaatggt gtctcatatc 8340  
 aatattggaa cagggtattga ctgcacgatt tgtgagcttg cggaaacaat agcaaaagtt 8400  
 gtagggtata aagggcatac tacgttcgat acaacaagc cogatggagc ccctcgaaaa 8460  
 ctacttgatg taacgcttct tcatcaacta ggttggaaac ataaaattac ccttcacaag 8520  
 ggtcttgaaa atacatacaa ctggtttctt gaaaaccaac ttcaatatcg ggggtaataa 8580  
 tgtttttaca ttcccaagac ttgccacaa ttgtaaggtc tactcctctt attctatatg 8640  
 atttgattgt ggaacacgag ttggcgcaaa ttttgctagg aaaacgaatc aaccgcccg 8700  
 cacagggtca ttggttcgtt cctggtggta ggtgttgtaa agatgaaaaa ttgcagacag 8760  
 ccttgaaacg attgacagaa attgaactag gaattcgttt gcctctctct gtgggtaagt 8820  
 ttatggttat ctggcagcac ttctacgaag acaatagatc ggggggagac ttttcaacgc 8880  
 attatatatt tatagcattc cttcttaaat tacaaccaa cattttgaaa ttaccgaagt 8940  
 cacaacataa tgcttattgc tggctatcgc gagcaaatg gataaatgat gacgatgtgc 9000  
 attataatgt tcgcgcatac tttaacaata aaacaaatga tgcgattggc tttagtaata 9060  
 aggatataat atgtctgatg cgcaataat tgcgttagtt atggccgggtg gtacaggcag 9120  
 tgcgtctttg ccactttctc gtgaactata tccaaagcag tttttacaa tcttgggtga 9180  
 taacaccttg ttacaaacga ctttgctacg actttcaggc ctatcatgtc aaaaaccatt 9240  
 agtgataaca aatgaacagc atcgctttgt tgtggctgaa cagttaaagg aaataataa 9300  
 attaaatggt aatatattc tagaaccatg cgggcgaaat actgcaccag caatagcgat 9360  
 atctcgcttt catcgcttaa aacgtaatcc tcaggaagat ccattgcttc tagttcttgc 9420  
 ggcagaccac gttatagcta aagaaagtgt tttctgtgat gctattaaaa atgcaactcc 9480  
 catcgctaac caaggtaaaa ttgtaacggt ttggaattata ccagaatatg ctgaaactgg 9540  
 ttatgggtat attgagagag gtgaactatc tgtaccgctt caaggcgatg aaaaatactg 9600  
 tttttattat gtaaaaatgt ttgtcgaaaa gcctaattct gaaacgcgag aattgtatat 9660  
 gactctctgt aatcacattt ggaatagttg aatattcatg tttaaggcat ctgtttatct 9720  
 tgaggaaatt agaaaaatta gacctgacat ttacaatggt tgtgaacagg ttgcctcatc 9780  
 ctctacattt gatctagatt ttattcgatt atcaaaagaa caatttcaag atcttctctg 9840  
 tgaatctatt gattttgctg taatggaaaa aacagaaaaa tgtgttgatt gcctgttga 9900  
 tattggttgg agtgacgttg gatcttgcca atcgttatgg gacattagtc taaaatcgaa 9960  
 aacaggagat gtatgtaaa ggtgatattt aacctatgat actaagaata attatactca 10020  
 ctctgagtca gcgttggtag ccgccattgg aattgaagat atggttatcg tgcaaacctc 10080  
 agatgccgtt cttgtgtcta aaaagagtga tgtacagcat gtaaaaaaaa tagtcgaaat 10140  
 gcttaattgt cagcaacgta cagagtatat tagtcatcgt gaagttttcc gaccatgggg 10200  
 aaaaattgat tgcattgacc aaggtgagcg atacaaaatc aagaaaaatc ttgtgaaacc 10260  
 tggtaggggg ctttctttaa ggaatgcatc ccatcgcttc gaacattgga tctgtcttcc 10320  
 tggtagcaga aaagtaaac ttggcgataa aactaaacta gtcaccgcaa atgaatcgat 10380  
 atacattccc cttggcgag cgtatagctc tgagaatacc ggcaatacc ctcttaactc 10440  
 tattgaagtc agttcagggg attatttggg agaggatgat attataagac agaaaagaac 10500  
 ttcaaacatc gaagtataac atatgaatc tttaacctgc tttaagacct atgatattcg 10560  
 cgggaaatta ggcgaagaac tgaatgaaga tattgcttgg cgcattgggc gtgcctatgc 10620  
 cgaatttctc aaacgaaaaa ccattgtttt agggcggtgat gtcgcgctca ccagcgaaag 10680  
 gttaaaactg gcgcttgca aaggtttaca ggatgcgggc gtcgatgtgc tggatatcgg 10740  
 tatgtccggc accgaagaga tctatttcgc caggttccat ctcgagtggt atggcgccat 10800  
 cgaagtacc cgcgcatac accgatgga ttacaacggc atgaagctgt tgcgcgaagg 10860  
 ggcctgcccg atcagcgggt ataccggact gcgcgatgtc cagcgtctgg cagaagccaa 10920  
 tgacttccct cctgtcgtat aaaccaaacg tggctcgcat cagcaaatca acttcggtga 10980  
 cgcttacgtt gatcacctgt tcggttatat caacgtcaaa accctcagc gtctcaagct 11040  
 ggtgatcaac tccgggaacg gcgcagcggt tccggtgggt gacgccattg aagcccgatt 11100  
 taaagccctc ggcgcaccgg tggaaataat caaagtacac aacacgccg acggcaattt 11160

ccccaacggt attcctaacc cgctgctgcc ggaatgccgc gacgacacc gtaatgcggt 11220  
 catcaaacac ggcgcggata tgggcattgc ctttgatggc gatattgacc gctgtttcct 11280  
 gtttgacgaa aaagggcagt ttatcgaggc ctactacatt gtcgcctgcg tggcagaagc 11340  
 gttcctcgaa aaaaatcccg gcgcgaagat catccacgat ccacgctctt cctggaaacac 11400  
 cgttgatgtg gtgactgccg caggcgccac cccggtaatg tcgaaaaccg gacacgcctt 11460  
 tattaaagaa cgtatgcgca aggaagacgc catctacggt ggcgaaatga gcgctcacca 11520  
 ttactccggt gatttcgctt actgcgacag cggcatgata ccgtggctgc tggcgcgga 11580  
 actggtgtgc ctgaaaggaa aaacgctggg cgaatgggtg cgcgaccgga tggcggcggt 11640  
 tcgcgcaagc ggtgagatca acagcaaac ggcgcaaccc gttgaggcaa ttaatcgcggt 11700  
 ggaacagcat tttagccgcg aggcgctggc ggtggatgcg accgatggca tcagcatgac 11760  
 ctttgcgcac tggcgcctta acctgcgctc ctccaacacc gaaccgggtg tgcggttgaa 11820  
 tgtggaatca cgcggtgatg taaagctaag gaaaaagaaa actaaagctc ttcttaaat 11880  
 gctaagtga tgatattta cattaatcat taagcgattt taagattata ttaaagtaat 11940  
 ttttgcggtg tatctatga atatgtgggc tttttatgt ataacgacta taccgcaact 12000  
 ttatctagga aaagattaat agaaataaag ttttgtactg accaatttgc attcacgctc 12060  
 acgattgaga cgctctcttg cttaagacat tttttcactg cttatgtaat acaaaatgtg 12120  
 cttatataaa aaaggagaac aaaaatggaac ttaaaataat tgagacaata gattttttat 12180  
 atccctgttt acgatattat agccaaagtt gtatcctgca tcagtcctgc aatatttcac 12240  
 gagtgtcttg ttaactgaat acatgtctgc cattttccag atgataacga cgtcatcgca 12300  
 attgcttgta aaacacttgc gcacacttat gacaagagtc gtgcgagagc agtggtttcat 12360  
 gtcattagtg cgtttcagca atgcacagtc tggctcctcg atagatcaag accggatgaga 12420  
 aaactaatgc gtctacagtt attcatgaac tttctaaaaa gatgggtatt aaaggaaaaa 12480  
 taatcataac tgatgcgatg gcttgccaga aagatattgc agagaagata taaaaacaga 12540  
 gatgtgatta tttattcgct gtaaaaggaa ataagagtcg gcttaataga gtctttgagg 12600  
 agatatattc cgtgaaagaa ttaataatc caaaacatga cagttacgca attagtgtaa 12660  
 agaggccagg cagagacgat gctcgtcttc atattgtttg agatgctcct gatgagctta 12720  
 ttgatttcac gtttgaatgg aaagggtgcg agaatttatg aatggcagtc cactttctct 12780  
 caataatgag agagcaaaag aaagaatccg aaatgacgat caaatattat attagatctg 12840  
 ctgctttaac cgcagagaag ttccgccacag taaatcgaaa tcactggcgc atggagaata 12900  
 agttgcacag tagcctgatg tggtaatgaa tgaatcgac tataataata gaaggcgagt 12960  
 tgcattcgaa tgattttcta gaatgcggca catcgctatt aatatctgac aatgataatg 13020  
 tattcaaggc aggattatca tgaagatgc gaaaagcagt catggacaga aacttctcat 13080  
 cgtcaggcat tgcagcgtgc gggctttcat aatcttgcat tggtttgat aagattattc 13140  
 tttggagatg gggaaatgaa ttgtatggt atttttggtg ctggaagtta tggtagagaa 13200  
 acaataccca ttctaaatca acaataaag caagaatgtg gttctgacta tgctctggtt 13260  
 tttgtggatg atgttttggc aggaagaaa gttaatggtt ttgaagtgc ttcaaccac 13320  
 tgctttctaa aagccctcta tttaaaaaag tattttaatg ttgctattgc taatgataag 13380  
 atacgacaga gagtgtctga gtcaatatta ttacacgggg ttgaaccaat aactataaaa 13440  
 catcaaaata gcgtgttcta tgatcatact atgataggta gtggcgctat tttctctccc 13500  
 ttgtttacaa tatctactaa tactcatata gggaggtttt ttcatgcaaa catatactca 13560  
 tacgttgcaac atgattgtca aataggagac tatgtttacat ttgctcctgg ggctaaatgt 13620  
 aatggatatg ttgttattga agacaatgca tatataggct cgggtgcgat aatgaagcag 13680  
 ggtgttccca atcgcacact tattattggc gcgggagcca ttataggat gggggctgtt 13740  
 gtcaactaaa gtgttctcgc cgggtataact gtgtgcgaaa atccagcaag agaaatgaaa 13800  
 agatcgccaa catctattta atgggaatgc gaaaacacgt tccaaatggg actaatgttt 13860  
 aaaaatatata taatttcgct aatttactaa attatggctt ctttttaagc tatcctttac 13920  
 ttagtattata ctgatcacgc atgaaatata taatactctg atacattttt atacgtttatt 13980  
 caagcgcgat attacgcgt aaccctgac aggagtaaac atag 14024

- 47 -

&lt;210&gt; 57

&lt;211&gt; 1758

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 57

```

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaataa tatcaacaag 60
aaccagtcctg cgctgtcgag ttctatcgag cgtctgtctt ctggcgttcg tatcaacagc 120
gcgaaggatg acgcccgcagg tcaggcgatt gctaaccgtt ttacttetaa cattaaggc 180
ctgactcagg cggcccgtaa cgccaacgac ggtatttctg ttgcgcagac caccgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtattctgt aactgacggt tcaggccact 300
acaggggacta actccgatc tgacctggac tccatccagg acgaaataca atctcgtctt 360
gatgaaattg accgcgtatc cggccagacc cagttcaacg cgtggaacgt gctggcgaaa 420
gacgggtcaa tgaataatca ggttggtgag aatgacggcg aaaccatcac gatcgacctg 480
aaaaaaatcg attctgatac tctgggtctg aatggcttta acgtaaatg taaaggtaact 540
attaccaaca aagctgcaac ggtaagtgat ttaacttctg ctggcgcgaa gttaaacacc 600
acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgcgtcattc 660
gataaattag ggaatggcga taaagtcaca gttggcgccg tagattatac ttacaacgct 720
aaatctgggt attttactac cactaaatct actgctggta cgggtgtaga cgccgcggcg 780
caggctgctg attcagcttc aaaacgtgat gcgttagctg ccacccttca tgctgatgtc 840
ggtaaatctg ttaattggtc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900
tcagcaggtg atatcaccat cggtggaagc caggcatacg tagacgatgc aggcacactg 960
acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgtctgc caagcagcg 1020
agcgaaggtg gtgacggtgc ctctctgaca ttcaatggca cagaatatac catcgcaaaa 1080
gcaactcctg gcacaaccac tccagtagct ccgttaatcc ctggtgggat tactttacag 1140
gctacagtga gtaaaagtgt agtattgagc gaaaccaaag cggctgcgcg gacatcttca 1200
attaccttta attccggtgt actgagcaaaa actattgggt ttaccgcggg tgaatccagt 1260
gatgctgcga agtcttatgt ggatgataaa ggtggtatca ctaacgttgc cgactataca 1320
gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgcgggata tgcctcagcg 1380
actgatacca ataaagatta tgctccagca attggtactg ctgtaaatgt gaactcccg 1440
ggtaaatca ctactgagac taccagtgtc ggttctgcaa cgaccaaccc gctgtgtgcc 1500
ctggcagcag caatcgacaa catcgacaaa ttccgttctt cctcggtgtc tatccgaac 1560
cgtctggatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc cgaagcgcag 1620
tcccgtatc aggacgcga ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatc 1680
attcagcagg ccggttaact cgtgctggca aaagctaacc aggtaccgca gcaggttctg 1740
tctctgctgc agggttaa

```

&lt;210&gt; 58

&lt;211&gt; 1758

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 58

```

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaataa tatcaacaag 60
aaccagtcctg cgctgtcgag ttctatcgag cgtctgtctt ctggcgttcg tatcaacagc 120
gcgaaggatg acgcagcggg tcaggcgatt gctaaccgtt ttacttetaa cattaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcgcagac caccgaaggc 240

```

gcgctgtccg aaatcaacaa caacttacag cgtattcgtg aactgacggt tcagggccact 300  
 acaggggacta actccgattc tgacctggac tccatccagg acgaaatcaa atctcgtctt 360  
 gatgaaattg accgcgtatc cggccagacc cagttcaacg gcgtgaacgt gctggcgaaa 420  
 gacgggtcaa tgaaaattca ggttggtgcy aatgacggcg aaaccatcac gatcgacctg 480  
 aaaaaaatcg attctgatac tctgggtctg aatggcttta acgtaaatgg taaaggctact 540  
 attaccaaca aagctgcaac ggtaagtgt ttaactctgt ctggcgcgaa gttaaaccacc 600  
 acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgcatte 660  
 gataaattag ggaatggcga taaagtcaca gttggcgggc tagattatata ttacaacgct 720  
 aaatctcgtg attttactac cactaaatct actgctggta cgggtgtaaa cgcccgggcg 780  
 caggctgctg attcagcttc aaaaacgtgt gcgttagctg ccacccttca tgctgatgtg 840  
 ggtaaatctg ttaattggctc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900  
 tcagcaggta atatacaccat cggtggaagc caggcatatc tagacgatgc aggcaacttg 960  
 acgataaca acgctggtag cgcagctaaa gctgatata aagcgtgctc caaagcagcg 1020  
 agcgaaggta gtgacgggtc ctctctgaca tccaatggca cagaatatac catcgcaaaa 1080  
 gcaactcctg cgacaaccac tccagtagct ccgttaatcc ctggtgggat tacttatcag 1140  
 gctacagtga ttaaagatgt agtattgagc gaaaccaaag cggctgcgcg gacatcttca 1200  
 attaccttta attccgggtg actgagcaaa actattgggt ttaccgcggtg tgaatccagt 1260  
 gatgctcgga agtcttatgt ggatgataaa ggtggtatca ctaacgttgc cgactatata 1320  
 gtctcttata cggttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380  
 actgatacca ataaagatta tgctccagca attggcactg ctgtaaatgt gaactccgct 1440  
 ggtaaaatca ctactgagac taccagtgtc ggtttctgaa cgaccaaccc gcttgctgcc 1500  
 ctggacgagc caatcagctc catcgacaaa ttcgttctt cctcggtgtc tatccgaac 1560  
 cgtctgatt ccgcggtcac caacctgaac aacaccaata ccaacctgtc cgaagcgagc 1620  
 tcccgtatc aggacgcgca ctatcgacc gaagtgtcaa acgtgtcgaa agcgcgagtc 1680  
 atccagcagc ccggttaactc cgtgctggca aaagctaacc aggtaccgca gcaggttctg 1740  
 tctctgctgc aggggttaa 1758

&lt;210&gt; 59

&lt;211&gt; 1758

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 59

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaata tatcaacaag 60  
 aaccagctcg cgctgtccgag ttctatcgat cgtctgtctt ctggcttgcg tattaacagc 120  
 gcgaaggatt acgcgcggcg tcaggcgatt gctaacgctt ttactcttaa cattaaaggc 180  
 ctgactcagg ctgcacgtaa cgccaacgac ggtattttctg ttgcacagac cactgaaggc 240  
 gcgctgtccg aaatcaacaa caacttacag cgtatccgtg agctgacggt tcaggtctct 300  
 accgggacta actctgattc ggaatctggac tccattcagc acgaaatcaa atcccgcttc 360  
 gagcaaatg accgcgtatc cggctcagacc cagttcaacg gcgtgaacgt actggcaaaa 420  
 gacgggttga tgaaaattca ggttggtgcy aatgacgggt aaactatcac tatcgacctg 480  
 aagaaaaatc attctgatac tctgggtctg aatggtttta acgtaaatgg taaaggctact 540  
 attaccaaca aagctgcaac ggtaagtgt ttaactctgt ctggcgcgaa gttaaaccacc 600  
 acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgcatte 660  
 gataaattag ggaatggcga taaagtcacc gttggcgggc tagattatata ttacaacgct 720  
 aaatctcgtg attttactac caccaaatct actgctgta cgggtgtgata cgcccgggcg 780  
 caggctactg attcagctaa aaaaacgtgt gcgttagctg ccacccttca tgctgatgtg 840  
 ggtaaatctg ttaattggctc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900

```

tcagcaggta atatcaccat cgggtggaagc caggcatatcg tagacgatgc aggcaacttg 960
acgactaaca acgctggttag cgcagctaaa gctgatataa aagcgcgtgct taaagccgcg 1020
agcgaaggta gtgacggtgc ctctctgaca ttcaattgca ctgaatatac tatcgcaaaa 1080
gcaactcctg cgacaacctc tccagtagct ccgttaatcc ctggtgggag tactttacag 1140
gtcacagtga gtaaaagatgt agtattgagc gaaaccaaag cggctgcgcg gacatcttca 1200
attaccttta attccgggtg actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
gatgctgcga agtcttatgt ggatgataaa ggtgggtatta ctaacgttgc cgactatata 1320
gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380
actgatacca ataaagatta tgctccagca attggtactg ctgtaaatgt gaactccgcg 1440
ggtaaaatca ctactgagac taccagtgtc ggttctgcaa cgaccaaccc gcttgctgcc 1500
ctggacgagc ctatcagctc catcgacaaa ttccgttctt ccttgggtgc tatccagaac 1560
cgctctgatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc tgaagcgagc 1620
tcccgattc aggacgccga ctatcgacc gaagtgtcca acatgtcgaa agcgcagatt 1680
atccagcagg ccggtaactc cgtgctggca aaagccaacc aggtaccgca gcaggttctg 1740
tctctgctgc aggggttaa 1758

```

&lt;210&gt; 60

&lt;211&gt; 1758

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 60

```

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtcgt cgctgtcgag ttctatcgag cgtctgtctt ctgggttgcg tattaacagc 120
gcgaaggatg acgcgcgagg tcaggcgatt gctaaccggt ttactttcaa cattaaggcg 180
tgactcagcg cgcccggtaa cgccaacgac ggtatttctg ttgcgcgagc caccgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtattctgt aactgacgtg agtcggcaact 300
acagggacta actccgattc tgacctggac tccatccagg acgaaatcaa atctcgtctt 360
gatgaaattg accgcgtatc cggccagacc cagttcaacg gcgtgaacgt gctggcgaaa 420
gacggttcaa tgaaaattca ggttgggtgc aatgacggcg aaaccatcac gatcgacctg 480
aaaaaaatcg attctgatac tctgggtctg aatggcttta acgtaaatgg taaagggtact 540
attaccaaca aagctgcaac ggtaagtgat ttaacttctg ctggcgcgaa gttaaaccac 600
acgacaggtc ttatgatctt gaaaaccgaa aataccttgt taactaccga tctgcattc 660
gataaattag ggaatggcga taaagtcaac gtggcgcgcg tagattatac ttacaacgct 720
aaatctgggt attttactac cactaaatct actgctggta cgggtgtaga cgccgcggcg 780
acggtctgct attcagcttc aaaaacgtgat gcgttagctg ccacccctca tgcgtgatgtg 840
ggtaaatctg ttaatggttc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900
tcagcaggta atatcaccat cgggtggaagc caggcatatcg tagacgatgc aggcaacttg 960
acgactaaca acgctggtag cgcagctaaa gctgatataa aagcgcgtgt caaagcagcg 1020
agcgaaggta gtgacggtgc ctctctgaca ttcaattgca cagaatatac catcgcaaaa 1080
gcaactcctg cgacaaccac tccagtagct ccgttaatcc ctggtgggag tactttacag 1140
gtacagtgta gtaaaagatgt agtattgagc gaaaccaaag cggctgcgcg gacatcttca 1200
attaccttta attccgggtg actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
gatgctgcga agtcttatgt ggatgataaa ggtgggtatca ctaacgttgc cgactatata 1320
gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380
actgatacca ataaagatta tgctccagca attggtactg ctgtaaatgt gaactccgcg 1440
ggtaaaatca ctactgagac taccagtgtc ggttctgcaa cgaccaaccc gcttgctgcc 1500
ctggacgagc caatcagctc catcgacaaa ttccgttctt ccttgggtgc tatccagaac 1560

```

```

cgtctggatt cgcgagtcac caacctgaac aacaccacta ccaacctgtc cgaagcgag 1620
tcccgtattc aggcgcgcga ctatgcgcacc gaagtgtcca acatgtcgaa agcgagatc 1680
attcagcagg ccggttaact cgtgctggca aaagctaacc aggtaccgca gcaggttctg 1740
tctctgctgc aggggttaa
1758

```

&lt;210&gt; 61

&lt;211&gt; 1758

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 61

```

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgtgtcgag ttctatcgag cgtctgtctt ctggcttgccg tattaacagc 120
gcgaaggatg agccgcgcag tcaggcgatt gctaaccgtt ttacttctaa cattaaggcg 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcgcgacac caccgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtattctgt aactgacggt tcaggccact 300
acagggacta ctccgattc tgacctggac tccatccagg acgaaataca atctcgtctt 360
gatgaaattg accgcgtatc cggccagacc cagttcaacg cgtggaacgt gctggcgaaa 420
gcaggttcaa tgaaaaatca ggttggtgcg aatgacggcg aaaccatcac gatcgacctg 480
aaaaaaatcg attctgatac tctgggtctg aatggcttta acgtaaatgg taaaggtaact 540
attaccaaca aagctgcaac ggtaagtgtt ttaactctg ctggcgcgaa gttaaacacc 600
acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgcattc 660
gataaattag ggaatggcga taaagtcaca gttggcgccg tagattatac ttacaacgct 720
aaatctgggt attttactac cactaaatct actgtgggta cgggtgtaga cgccgcggcg 780
caggctctgt attcagcttc aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840
ggtaaatctg ttaatgggtt ttacaccaca aaagatggta ctgtttctt cgaaacggat 900
tcagcaggtt atatcaccat cgttggaagc caggcatacg tagacgatgc aggcgaactg 960
acgactaaca acgtggttag cgcagctaaa gctgatatga aagcgtctgt caaagcagcg 1020
agcgaaggta gtgacggtgc ctctctgaca ttcaatggca cagaataatc catcgcaaaa 1080
gcaactcctg cgacaaccac tccagtagct ccggttaatc ctgggtggag tacttatcag 1140
gtacagtgta gtaaatgtgt agtattgagc gaaaccaaag cggctgccgc gacattctca 1200
attaccttta attccggtgt actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
gatgctgcga agtcttatgt ggatgataaa ggtggtatca ctaacgttgc cgactataca 1320
gtctcttaca cgtttaacaa ggataacggc tctgtgactg ttgcccggga tgcttcagcg 1380
actgatacca ataaagatta tgcctccagca attggcactg ctgtaaatgt gaactccgcg 1440
ggtaaaatca ctactgagac taccagtgtc ggttctgcaa cgaccaaccc gcttgctgcc 1500
ctggacgacg caatcagcto catcgacaaa ttccgttctt cctcgggtgc tatccgaac 1560
cgtctggatt cgcggttcac caacctgaac aacaccacta ccaacctgtc cgaagcgag 1620
tcccgatttc aggcgcgcga ctatgcgcacc gaagtgtcca acatgtcgaa agcgagatc 1680
atccagcagg ccggttaact cgtgctggca aaagctaacc aggtaccgca gcaggttctg 1740
tctctgctgc aggggttaa
1758

```

&lt;210&gt; 62

&lt;211&gt; 1758

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 62



atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaataa tatcaacaag 60  
 aaccagtcctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgccg tattaacagc 120  
 gcgaaggatg acgcgcgcggg tcaggcgatt gctaaccgtt ttactttctaa cattaaaggc 180  
 ctgactcagg ctgcacgttaa cgccaacgac ggtattttctg ttgcacagac cactgaaggc 240  
 gcgctgtccg aaatcaacaa caacttacag cgtatccgtg agctgacggt tcaggcttct 300  
 accgggacta actctgattc ggatctggac tccattcagg acgaaatcaa atccgtctc 360  
 gacgaattg accgcgtatc cggctcagacc cagttcaacg cgtgaaactg actggcaaaa 420  
 gacggttcga tgaaaattca ggttggtgag aatgacggg aaactatcac tatcgacctc 480  
 aagaaaatcg attctgatac tctgggtctg aatgggttta acgtaaatgg taaaggctact 540  
 attaccaaca aagctgcaac ggtaagtgt ttaactctg ctggcgcgaa gttaaacacc 600  
 acgacaggtc tttatgatct gaaaaccgaa aatacctgt taactaccga tctgctattc 660  
 gataaattag ggaattggca taaagtcacc gttggcgccg tagattatac ttacaacgct 720  
 aaatctggtg attttactac caccaaatct actgctggta cgggtgtaga cgcgcgcggc 780  
 caggctactg attcagctaa aaaacgtgat gcgttagctg ccacccttca tgctgtatgt 840  
 ggtaaatctg ttaattggtc ttacaccaca aaagatggta ctgtttctt cgaacaggat 900  
 tcagcaggta atatcaccat cgggtggaagc caggcatatc tagacgatg aggcaactg 960  
 acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgtctgt taaagccgcg 1020  
 agcgaaggta gtgacgggtc ctctctgaca ttcaatggca ctgaatatac tatcgcaaaa 1080  
 gcaactctg cgacaacctc tccagtagct ccgttaatcc ctggtgggat ttcttatcag 1140  
 gctacagatg gtaaaagtgt agtattgagc gaaaccaaag cggctgccgc gacatcttca 1200  
 attaccttta attccggtgt actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260  
 gatgctcgca agtcttatgt ggatgataaa ggtggtatta ctaacgttgc cgactataca 1320  
 gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttggccggga tgcttcagcg 1380  
 actgatacca ataaagatta tgctccagca attggtactg ctgtaaatgt gaactccgcg 1440  
 ggtaaaatca ctactgagac tacagtgctt ggtttctgaa cgaccaaccc gcttctgccc 1500  
 ctggacgacg ctatcagctc catcgacaaa ttccgttctt ccttgggtgc tatccagaac 1560  
 cgtctggatt ccgcagtcga caacctgaac aacaccacta ccaacctgtc tgaagctgac 1620  
 tcccgtattc aggcacgcga ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatt 1680  
 atccagcagg ccggttaact cgtgctggca aaagccaacc aggtaccgca gcaggttctg 1740  
 tctctgtgc agggttaa 1758

&lt;210&gt; 63

&lt;211&gt; 1758

&lt;212&gt; DNA

<213> *Escherichia coli*

&lt;400&gt; 63

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaataa tatcaacaag 60  
 aaccagtcctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgccg tattaacagc 120  
 gcgaaggatg acgcgcgcagg tcaggcgatt gctaaccgtt ttactttctaa cattaaaggc 180  
 ctgactcagg cgcccgctaa cgccaacgac ggtattttctg ttgcacagac caccgaaggc 240  
 gcgctgtccg aaatcaacaa caacttacag cgtatccgtg aactgacggt tcaggcttct 300  
 acaggggacta actccgattc tgacctggac tccatccagg acgaaatcaa atctcgttct 360  
 gatgaaattg accgcgtatc cggccagacc cagttcaacg cgtgaaactg gctggcgaaa 420  
 gacggttcga tgaaaattca ggttggtgag aatgacggcg aaacctatcac gatcgacctg 480  
 aaaaaaatcg attctgatac tctgggtctg aatggcttta acgtaaatgg taaaggctact 540  
 attaccaaca aagctgcaac ggtaagtgt ttaactctg ctggcgcgaa gttaaacacc 600  
 acgacaggtc tttatgatct gaaaaccgaa aatacctgt taactaccga tctgctattc 660

gataaattag ggaattggcga taaagtcaca gtggcgccg tagattatac ttacaacgct 720  
 aaatctgggtg attttactac cactaaatct actgctggta cgggtgtaga cgccgcggcg 780  
 caggctgctg attcagcttc aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840  
 ggtaaatctg ttaattgggtc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900  
 tcagcaggtg atataccat cgggtggaagc caggcatatg tagacgatgc aggcaacttg 960  
 acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgctgct caaagcagcg 1020  
 agcgaaggtg gtgacgggtc ctctctgaca ttcaatggca cagaatatac catcgcaaaa 1080  
 gcaactctcg cgacaaccac tccagtagct ccgttaatcc ctgggtgggt tactatcag 1140  
 gctacagtga gtaaagatgt agtattgagc gaacccaaag cggtctgcgc gacatctcca 1200  
 attacccttta attccgggtg actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260  
 gatgctgcga agtcttatgt ggatgataaa ggtgggtatca ctaacgttgc cgactataca 1320  
 gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380  
 actgatacca ataaagatta tgctccagca attggtactg ctgtaaatgt gaactccgcg 1440  
 ggtaaaatca ctactgagac taccagtgtc ggttctgcaa cgaccaaccg gctgtctgcc 1500  
 ctggacgacg caatcagctc catcgacaaa ttccgttctt ccttgggtgc tatccagaac 1560  
 cgtctggatt ccgcagctac caacctgaac aacaccacta ccaacctgtc cgaagcgcag 1620  
 tcccgatttc aggcgcgcga ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatc 1680  
 attcagcagg ccggttaactc cgtgctggca aaagctaacc aggtaccgca gcaggttctg 1740  
 tctctgctgc aggggttaa 1758

&lt;210&gt; 64

&lt;211&gt; 1758

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 64

atggcacaag tcattaatac caacagctc tcgctgatca ctcaaaataa tatcaacaag 60  
 aaccagttct cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120  
 gcaaggatg acgcgcggcg tcaggcgatt gctaaccggt ttacttctaa cattaaggcg 180  
 ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcacagac caccgaagcg 240  
 gcgctgtctg aaatacaaca caacttacag cgtatccgtg agctgacggt tcaggcttct 300  
 accggaacta actctgattc ggatctggac tccattcagg acgaaatcaa atccggtctt 360  
 gatgaaattg accgcgtatc cggccagacc cagtccaacg gcgtgaacct actggcaaaa 420  
 gacggttcga tgaaaattca ggttgggtcg aatgacgggt aaactatcac taacgacctg 480  
 aagaaaatcg attctgatac tctgggtctg aatggtttta acgtaaatgg taaaggctact 540  
 attaccaaca aagctgcaac ggtaagtgtg taaactctg ctggcgcgaa gttaaacacc 600  
 acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgacttc 660  
 gataaattag ggaattggcga taaagtcacc gttgcggcg tagattatac ttacaacgct 720  
 aaatctgggt attttactac caccaaatct actgctggta cgggtgtaga cgccgcggcg 780  
 caggctactg attcagctaa aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840  
 ggtaaatctg ttaattgggtc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900  
 tcagcaggtg atataccat cgggtggaagc caggcatatg tagacgatgc aggcaacttg 960  
 acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgctgct taagccgcg 1020  
 agcgaaggtg gtgacgggtc ttctctgaca ttcaatggca ctgaatatac tatcgcaaaa 1080  
 gcaactctcg cgacaaccac tccagtagct ccgttaatcc ctgggtgggt tactatcag 1140  
 gctacagtga gtaaagatgt agtattgagc gaacccaaag cggtctgcgc gacatctcca 1200  
 attacccttta attccgggtg actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260  
 gatgctgcga agtcttatgt ggatgataaa ggtgggtatta ctaacgttgc cgactataca 1320

tctctttaca gcgttaacaa ggataacggc tctgtgactg ttgcgggta tgcttcagcg 1380  
 actgatacca ataagatta tgctccagca attggtactg ctgtaaatgt gaactccgcg 1440  
 ggtaaaatca ctactgagac taccagtgtt ggttctgcaa cgaccaaccc gcttgcgtcc 1500  
 ctggacgacg ctatcagctc catcgacaaa ttccgttctt ccttgggtgc tatccagaac 1560  
 cgtctggatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc tgaagcgcag 1620  
 tcccgtattc aggacgcgca ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatt 1680  
 atccagcagg ccggttaact cgtgctggca aaagccaacc aggtaccgca gcaggttctg 1740  
 tctctgctgc agggtttaa 1758

&lt;210&gt; 65

&lt;211&gt; 1758

&lt;212&gt; DNA

<213> *Escherichia coli*

&lt;400&gt; 65

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaaata tatcaacaag 60  
 aaccagctgc cgtgtctgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120  
 gcgaaggatg acgcccgggg tcaggcgatt gtaaccggtt ttacttctaa cattaaggcg 180  
 ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcacagac cactgaaggc 240  
 gcgctgtccg aaatcaacaa caacttacag cgtatccgtg agctgacggt tcaggettct 300  
 accgggacta actctgattc ggatctggac tccattcagg acgaaatcaa atccggtctc 360  
 gacgaaatg acgcggtatc cggtcagacc cagttcaacg cgtggaacgt actggcaaaa 420  
 gacggtctga tgaaaattca ggttggtgag aatgacgggt aaactatcac tatcgacctg 480  
 aagaaaatcg attctgatac tctgggtctg aatggtttta acgtaaatgg taaaggatct 540  
 attaccaaca aagctgcaac ggtaagtgat ttaacttctg ctggcgcgaa gttaaacacc 600  
 acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgattcc 660  
 gataaattag ggaatggcga taaagtcacc gtggcgcg tagattatac ttacaacgct 720  
 aaatctgggt attttactac caccaaatct actgctggta cgggtgtaga cgccgcggcg 780  
 caggctactg attcagctaa aaaacgtgat gcgttagctg ccacccctta tgctgatgtg 840  
 ggtaaatctg ttaatggttc ttaccaccac aaagatggta ctgattcttt cgaacggat 900  
 tcagcaggta atatcccat cggtggaagc caggcatagc tagacgatgc aggcacttg 960  
 acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgtgct taaagccgcg 1020  
 agcgaaagga gtgacgggtc ctctctgaca ttcaatggca ctgaatatac tatcgcaaaa 1080  
 gcaactcctg cgacaacctc tccagtagct ccgttaatcc ctgggtggat ttcttatcag 1140  
 gctacagtga gtaaatagtg agtattgagc gaaacaaaag cggtgcgcgc gacatcttca 1200  
 attaccttta attccgggtg actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260  
 gatgctgcga agtcttatgt ggatgataaa ggtggtatta ctaacgttgc cgactataca 1320  
 gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgcgggta tgcttcagcg 1380  
 actgatacca ataaagatt ta tgctccagca attggtactg ctgtaaatgt gaactccgcg 1440  
 ggtaaaatca ctactgagac taccagtgtt ggttctgcaa cgaccaaccc gcttgcgtcc 1500  
 ctggacgacg ctatcagctc catcgacaaa ttccgttctt ccttgggtgc tatccagaac 1560  
 cgtctggatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc tgaagcgcag 1620  
 tcccgtattc aggacgcgca ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatt 1680  
 atccagcagg ccggttaact cgtgctggca aaagccaacc aggtaccgca gcaggttctg 1740  
 tctctgctgc agggtttaa 1758

&lt;210&gt; 66

&lt;211&gt; 1758

&lt;212&gt; DNA

<213> *Escherichia coli*

&lt;400&gt; 66

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60  
 aaccagctcg cgtcgtcgag ttctatcgag cgtcgtcttt ttggtcttgcg tattaacacg 120  
 gcgaaggatg acgcccgggg tcaggcgatt gctaaccggt ttactttctaa catcaaaggc 180  
 ctgactcagg ctgcacgtaa cgccaacgac ggtattttctg ttgcacagac cactgaaggc 240  
 gcgctgtccg aaatcaacaa caacttacag cgtatccgtg agctgacggt tcaggcttct 300  
 accgggatac actctgattc ggtctggac tcattcagg acgaaatcaa atcccgctc 360  
 gacgaaattg accgcgtatc cggctcagacc cagttcaacg gcgtgaacgt actggcaaaa 420  
 gacgggtcga tgaaaattca ggtagggtgcg aacgacggcc agactatcac tattgatctg 480  
 aagaaaattg actctgatac gctggggctg aatgggttta acgtgaatgg ttccgggtacg 540  
 atagccaata aagcggcgac cattagcgac ctgacagcag cgaaaatgga tgctgcaact 600  
 aatactataa ctacaacaaa taatgcgctg actgcatcaa aggcccttga tcaactgaaa 660  
 gatgggtgac ctgttactat caaagcagat gcagctcaaa ctgccacggt ctatacatac 720  
 aatgcactctg ctggtaactt ctcatctcagt aatgtatcga ataactctc agcaaaagca 780  
 ggtgatgtag cagctagcct tctcccgccg gctgggcaaa ctgctagtgg tgtttacaaa 840  
 gcgcaagcgt gtgaagttaa cttttagtgg gatcgcaatg gtaaaattac aatcgaggga 900  
 caggaagcct atttaactag tgatggtaac ttaactacaa acgatgctgg tggctgcgact 960  
 gcggctacgc ttgatgggtt attcaagaaa gctgggtgat gtcaatcaat cgggtttaat 1020  
 aagactgcac cagtcacgat ggggggaaca actataact ttaaaacggg tgctgatgct 1080  
 ggtgctgcaa ctgctaacgc aggggtatcg ttactgata cagctagcaa agaaacgct 1140  
 ttaataaag tggtctacgc taaacaagc acagcagttg cagctaacgg tgatacatcc 1200  
 gcaacaatta cctataaact tggcgcttcc agctatcagg cggtatttgc cgcagggtgac 1260  
 ggtactgcta gcgcaaaata tgccgataat actgacgttt ctaatgcaac agcaacatac 1320  
 acagatgctg attggtgaaat gactacaatt ggttcataca ccacgaagta ttcaatcgat 1380  
 gctaaacaac gcaaggtaac tgttgattct ggaactggtt cgggtaataa tgcgccgaaa 1440  
 gtcggggctg aagtatatgt tagtgctaatt ggtactttta caacagatgc aactagcgaa 1500  
 ggacagatga caaaagatcc actgaaagct ctggatgaag ctatcagctc catcgacaaa 1560  
 ttccgttcat cctcgggggc tatccaaaac cggttggtat ccgcccgtcac caacctgaac 1620  
 aacaccacta ccaacctgtc tgaagcgcag tcccgatttc aggcagccga ctatgcgacc 1680  
 gaagtgtcca acatgtcgaa agcgcagatt atccagcagg ccggttaactc cgtgctggga 1740  
 aaagccaacc aggtaccgca gcaggttctg tctctactgc aggggttaa 1788

&lt;210&gt; 67

&lt;211&gt; 1398

&lt;212&gt; DNA

<213> *Escherichia coli*

&lt;400&gt; 67

aacaaatctc agtcttctct tagctctgct attgagcgtc tgtcttcttg tctgcgtatt 60  
 aacagcgcaa aagacgatgc agcaggtcag gcgattgcta accgttttac gccaaatatt 120  
 aaaggtctga cccagctctc ccgtaacgca aatgatggta ttctgtttgc gcagaccact 180  
 gaaggtgcgc tgaatgaaat taacaacaac ctgcagcgta ttctgtaact ttctgttcag 240  
 gcaactaacy gtacttaactc tgacagtgcac ctgacctcca tccagtcoga aatccagcag 300  
 cgtctgagtg aaattgacgg tgtttctggg cagactcagt ttaacggcgt taaagtgctg 360  
 gcttctgatc aggatgatgc tattcagggt ggtgcaaacg accggcgaac aattactatt 420

```

aaactgcagg aaattaattc cgacacactg ggattatctg gttttggtat taaagatcct 480
actaaattaa aagccgcaac ggctgaaaca acctattttg gatcgacagt taagcttgct 540
cagcgtataa cacttgatgc agatattaca gctacagtta aaggcactac gactccgggc 600
caacgtgacg gtaaatattat gtctgatgct aacggtaagt tgcacgttaa agttgccggg 660
tcagataaac ccgctgaaaa tggttattat gaagttactg tggaggatga tccgacatct 720
cttgatgcag gtaagctgaa gctgggggct ctacggggta cccagcctca agctgggtaat 780
ttaaaggaa gtcacaacgg gaaagggaa ggggctattg atgttcagtt gggtagctgat 840
accgcaacgg ctctctatcac aggtgcaaaa ctctttaagt tagaagacgc caatggcaaa 900
gatactggtt catttgcgtt gatttggtgat gacggtaaac agtatgcagc gaatgttgat 960
cagaaaaacg gagcagtttc cgttaaaaca atgtcttaca ctgatgctga cgggtgtcaaa 1020
cacgacaatg ttaaagttga actgggtgga agcgatggca aaaccgaagt tgtaactgca 1080
accgatggca aaacttacag tgtagtgat ttacaaggta agagcctgaa aactgattct 1140
attgcagcaa tttctacga gaaaacagaa gatcctttgg ctgctatcga taagcactg 1200
tctcaggttg actcgttgcg tcttaacctg ggtgcaattc aaaaactgtt cagacttgcc 1260
atcaccaacc ttggcaaac cgtaaacaac ctgtcttctg cccgtagccg tatcgaagat 1320
gctgactacg cgaccgaagt gtctaacatg tctcgtgcgc agatcctgca acaagcgggt 1380
acctctgttc tggcgag 1398

```

&lt;210&gt; 68

&lt;211&gt; 1479

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;400&gt; 68

```

aacaaatctc agtcttctct gagctcggc attgaacgtc tctcttctgg cctgcgtatt 60
aacagtgtca aagatgacgc agcaggtcag gcgattgtca accgttttca gcgcaaatatt 120
aaaggtctga tgaaggtctc cgttaacgcg aatgatggta tttctgtttg cgagaccact 180
gaaggtgcgc tttctgaaat caacaataac ttacagcgta ttctgtaatt gtcagtacag 240
gccactaatg gtacaaactc tgactccgac ctgaattcaa ttcaggatga aattacacaa 300
cgctcttagt aaattgactg tgtttctaac cagacacaaat ttaattggtgt aaaagtctcg 360
gtctctgacg agactatgaa aattcaagta ggtgcgaacg atggtgaaac cattgagatt 420
gcctctgata aaattgatgc taaaaccttg gggcttgata accttagcgt agcaccagga 480
aaagttccaa tgcctctcgc ggttgcaact aagagcgaa cgcctcctga cttaaactaa 540
gtaaatgcaa ctgatggtag tgtgggaggt gctaaagcat tcggtagcaa ttataaaaat 600
gctgatgttg aaacttattt tggtaaccgt aatgtacaag atacaaggga tacaactgat 660
gcgaccggta ctgcaggaa aaaagtttat caagtacagg tggaagggca gacttatttt 720
gttggtcaag ataataatc caacacgaac ggttttcat tattgaaaca aaactctaca 780
ggttatgaaa aagttcaggt ggggtgtaag gatgttcagt tagcaaaact tgggtgtcgt 840
gtaactgcat ttgttgaga taatggttct gccacatcag ttgatttagc tgcgggttaa 900
atgggttaag cattagctta taatgatgca ccaatgtctg tttatttttg gggaaaaaac 960
ctagatgtcc accaagtaca agatacccaa ggaatcctg tacctaattc atttgtctct 1020
aaaaacatcag acggcaacta cattgcagta aatgtagatg ccgctacagg taacacgtct 1080
gttattactg atcctaattg taaggcagtt gaatgggcag taaaaaatga tggttctgca 1140
caggcaatta tgcgtgaaga tgataagggt tatcacgcca atatcacgaa taagaacgga 1200
accaaagggt ctgaactcag tgcctcagat ttgaaagcct tgcaaccac aaatccatta 1260
tccacattag acgaagcttt ggcaaaagtt gataagttgc gactgtcttt ggggtgcagta 1320
caaaacgctt tcgactctgc catcaccaac cttggcaaca ccgtaaacaa cctgtcttct 1380
gcccgtaggc gtatagaaga tgctgactac gcaaccgaag tgtctaactat gtctcgtgcg 1440

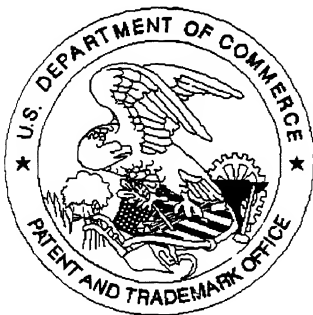
```

cagatcctgc aacaagcggg tacctctgtt ctggcacag

1479

(10)  
(11)  
(12)  
(13)  
(14)  
(15)  
(16)  
(17)  
(18)  
(19)  
(20)  
(21)  
(22)  
(23)  
(24)  
(25)  
(26)  
(27)  
(28)  
(29)  
(30)  
(31)  
(32)  
(33)  
(34)  
(35)  
(36)  
(37)  
(38)  
(39)  
(40)  
(41)  
(42)  
(43)  
(44)  
(45)  
(46)  
(47)  
(48)  
(49)  
(50)  
(51)  
(52)  
(53)  
(54)  
(55)  
(56)  
(57)  
(58)  
(59)  
(60)  
(61)  
(62)  
(63)  
(64)  
(65)  
(66)  
(67)  
(68)  
(69)  
(70)  
(71)  
(72)  
(73)  
(74)  
(75)  
(76)  
(77)  
(78)  
(79)  
(80)  
(81)  
(82)  
(83)  
(84)  
(85)  
(86)  
(87)  
(88)  
(89)  
(90)  
(91)  
(92)  
(93)  
(94)  
(95)  
(96)  
(97)  
(98)  
(99)  
(100)

United States Patent & Trademark Office  
Office of Initial Patent Examination – Scanning Division



SCANNED # 24

Application deficiencies found during scanning:

☒ Page(s) 63/96 of Drawings <sup>was</sup> ~~were~~ not present  
for scanning. (Document title)

☒ Page(s) 7/7 of Sequences were not present  
for scanning. (Document title)

☐ Scanned copy is best available.